UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE MATEMÁTICA

Técnicas de Álgebra Multilinear na Resolução de Problemas de Teoria Aditiva

Petra Alexandra Morbey Ferro Gaspar Pacheco

Mestrado em Matemática

Dissertação orientada por:
Professor Doutor Pedro J. Freitas

2019

Agradecimentos

O apoio incondicional de todos os que me acompanharam nesta grande caminhada foi, sem dúvida, o que tornou a conclusão desta dissertação possível! Sou profundamente grata por ter encontrado todos os demais, que me fortaleceram e me permitiram ultrapassar todos os desafios que esta dissertação me apresentou. As palavras nunca serão suficientes para agradecer todo o apoio que me deram, mas nem por isso deixarei de o tentar fazer.

Um enorme agradecimento aos meus amigos e colegas da licenciatura, foi um prazer descobrir Matemática convosco! Será um prazer continuar a fazê-lo! Muito obrigada por todos estes anos de companheirismo, simplesmente não há palavras para vocês!

Aos amigos e colegas que trago desde muito nova por sempre acreditarem em mim, por sempre me relembrarem que há motivos para continuar e que nenhum desafio é grande o suficiente que não seja superável!

Ao Jorge por todo o esforço em estar sempre presente! Por toda a compreensão, ajuda, paciência, por ser um porto seguro nos momentos menos bons e por sempre me encorajar a continuar! Obrigada por me tornares mais forte!

À Professora Maria Manuel por me ter introduzido, de forma tão entusiasmante e contagiante, à Álgebra Multilinear e pelo seu enorme apoio durante a adpatação ao mestrado. Um enorme agradecimento também à Professora Gracinda Cunha por todo o apoio durante o meu percurso académico, desde a licenciatura até ao fim do mestrado. Agradecer a ambas estas professoras por me mostrarem o melhor lado deste mestrado e por me inspirarem e estimularem a ser melhor.

À Professora Sónia Carvalho cujo brilhante percurso académico desde cedo me inspirou e ainda pelos seus comentários e correções que permitiram melhorar esta dissertação.

Um especial agradecimento ao meu orientador, o Professor Pedro Freitas, por todo o apoio incondicional que me deu desde o início deste Mestrado, que me acompanhou não só nos desafios académicos como nos desafios pessoais. Pela sua compreensão nos sucessivos atrasos da redação desta tese devidos à minha contínua necessidade de ocupar todo o meu tempo com diversos “projetos", pela sua constante disponibilidade, pela sua preocupação, pela sua prontidão em rever as sucessivas versões da tese e pela sua paciência. Obrigada por ser um Professor cuja paixão por Matemática contagia todos em seu redor! Obrigada por me mostrar as mais pequenas subtilezas da Matemática, obrigada por me orientar durante estes três anos! Simplesmente obrigada!

E por fim, e de todos o mais importante, um enorme agradecimento à minha família! Em especial aos meus pais que desde sempre me motivaram para alcançar todos os meus objetivos, que sempre me apoiaram em todos os momentos, a quem eu devo todas as oportunidades que hoje se me apresentam e a quem colocou sempre as minhas necessidades à frente das suas próprias! Foram, são e sempre serão o meu maior apoio e o meu maior motivo de orgulho!

Obrigada a todos por acreditarem em mim! Muito obrigada!

Resumo

Nesta dissertação expõe-se a aplicação de técnicas de Álgebra Multilinear no estudo de vários problemas de Teoria Aditiva. Os desenvolvimentos possibilitados por esta nova abordagem culminaram com a prova da conjetura de Erdős e Heilbronn demonstrando que

$$|A + A| \geqslant \min\{2|A| - 3, p\},$$

onde A é um subconjunto não vazio de um corpo com característica p, em relação ao qual o conjunto soma, apenas inclui somas de elementos distintos.

São ainda estudadas as principais diferenças e simplificações possibilitadas pela aplicação das referidas técnicas, através da exposição das demonstrações originais e das demonstrações simplificadas pela Álgebra Multilinear, do Teorema de Cauhcy-Davenport.

Os resultados apresentados devem-se principalmente ao trabalho realizado pelos Professores J. A. Dias da Silva e Yahya Hamidoune, em [6] e [7].

Palavras-chave: Teorema Cauchy-Davenport, tensores antissimétricos, Espaço de Grassmann, Conjetura de Erdős e Heilbronn, Método Polinomial.

Abstract

In this thesis, we study the application of procedures from Multilinear Algebra to multiple problems in Additive Number Theory. The breakthroughs achieved through this new approach allowed to prove the Erdős and Heilbronn's conjecture, showing that

$$|A + A| \geqslant \min\{2|A| - 3, p\},$$

where A is a nonempty subset of a field with characteristic p, whereas the sum set only includes the addition of distinct elements.

We will also study the main differences and simplifications made possible by applying such techniques mentioned, by presenting the original proofs as well as the simplified proofs from Multilinear Algebra, of the Cauchy-Davenport Theorem.

The main results presented in this thesis were possible due to the works of the Professors J.A. Dias da Silva and Yahya Hamidoune, in [6] and [7].

Keywords: Cauchy-Davenport Theorem, skew-symetric tensors, Erdős and Heilbronn's Conjecture, Polynomial Method.

Notações

- Dado $b \in \mathbb{R}$, escreve-se $\lfloor b \rfloor$ para denotar o maior inteiro menor ou igual a b.
- Dado $b \in \mathbb{R}$, escreve-se $\lceil b \rceil$ para denotar o menor inteiro maior ou igual a b.
- A cardinalidade de um dado conjunto A denota-se por $|A|$.
- Seja V um espaço vetorial sobre um corpo $\mathbb{F}$. A dimensão de V denota-se por $\dim V$.
- O determinante de uma matriz A representa-se por $\det(A)$.
- M_n representa o corpo das matrizes quadradas de ordem n.
- Dado um polinómio f, o grau de f é $\deg f$.
- Dado um espaço vetorial V sobre um corpo $\mathbb{F}$, id_V denota a aplicação identidade

$$\mathrm{id}_V : V \to V, \text{ tal que } \forall v \in V, \mathrm{id}(v) = v.$$

- Dada uma matriz A, denota-se por A^{-1} a matriz inversa de A, caso exista.
- Dada uma matriz A, denota-se por A^* a matriz transconjungada da matriz A.
- Dado um espaço vetorial V sobre um corpo $\mathbb{F}$, $(\vec{0})$ denota o vetor nulo do espaço V.
- Dados espaços vetoriais $V_1, \ldots, V_m, W$ sobre um corpo $\mathbb{F}$, o espaço vetorial das aplicações multilineares de $V_1 \times \ldots \times V_m$ sobre W denota-se por $L(V_1, \ldots, V_m; W)$.

  Sempre que $V_1 = \ldots = V_m$, o espaço anteriormente definido pode ser denotado por $L^m(V; W)$.
- O m-ésimo produto cartesiano de um espaço vetorial V, sobre um corpo $\mathbb{F}$, representa-se por $\times^m V$.
- Dados espaços vetoriais V e W sobre um corpo $\mathbb{F}$, o conjunto $L^m_A(V; W)$ denota o espaço vetorial das aplicações multilineares antissimétricas de $\times^m V$ para W.
- Dadas duas funções f e g, denota-se a sua composição por $f \circ g$, sempre que esta estiver bem definida.
- Dada uma função f, $\mathrm{Im}\ f$ denota o seu conjunto imagem.

- Sejam $V_1, \ldots, V_m$ espaços vetoriais sobre um corpo $\mathbb{F}$ de dimensão $n_1, \ldots, n_m$, respetivamente. Define-se o conjunto $\Gamma(n_1, \ldots, n_m)$ como

$$\Gamma(n_1, \ldots, n_m) := \{\alpha : \{1, \ldots, m\} \to \mathbb{N} : 1 \leqslant \alpha(i) \leqslant n_i,\ i = 1, \ldots, m\};$$

  O conjunto $G_{m,n}$ define-se como

$$\{\alpha \in \Gamma(m, n) : \alpha \text{ é crescente em sentido lato}\};$$

  E ainda o conjunto $Q_{m,n}$, que se define como

$$\{\alpha \in \Gamma(m, n) : \alpha \text{ é estritamente crescente}\}.$$

- Dada uma matriz $A \in \mathbb{M}_{m\times n}$, $\alpha = (\alpha_1, \ldots, \alpha_p) \in \Gamma(p, m)$ e $\beta = (\beta_1, \ldots, \beta_q) \in \Gamma(q, n)$, então $A[\alpha|\beta]$ é a matriz de $M_{p\times q}$, cuja entrada (i, j) é definida por $\left(a_{\alpha_i\beta_j}\right)$.
- Dado uma matriz $A \in \mathbb{M}_n$, o seu espectro denota-se por $\sigma(A)$.
- Dado um inteiro m, S_m representa o grupo simétrico de ordem m.
- Dada uma permutação $\sigma \in S_m$, $\operatorname{sgn}(\sigma)$ denota o sinal da permutação σ. Tem-se que $\operatorname{sgn}(\sigma) = 1$ se a permutação for par e $\operatorname{sgn}(\sigma) = -1$ se a permutação for ímpar.
- Dada uma partição λ, ξ_λ é o carácter irredutível associado a λ.
- Dados dois conjuntos A e B, a sua união disjunta denota-se por $A \dot{\cup} B$.
- Dados naturais m e q, $\binom{m}{q}$ denota o coeficiente binomial de m por q.
- Dados inteiros n e m, escreve-se $n \mid m$ sempre que o resto da divisão inteira de m por n seja nulo, caso contrário $n \nmid m$.
- Dados inteiros n, m e q diz-se que $n \equiv m (\mathrm{mod}\ p)$ se os restos das divisões inteiras de n por p e m por p coincidirem; caso contrário escreve-se $n \not\equiv m$.
- Dado um número inteiro n, denota-se $|n|$ o seu valor absoluto.
- Dados inteiros i e j, denota-se o delta de Kronecker por $\delta_{i,j}$.
- Dado um operador f num dado espaço vetorial, P_f denota o polinómio mínimo do operador.
- Dado um operador linear f num dado espaço vetorial, Df denota a derivada de Grassmann.
- O conjunto P_s denota o conjunto de todas as partições de comprimento no máximo s.
- O conjunto $P_{k,s}$ denota o conjunto das partições de grau k e comprimento no máximo s.

- Dados v e w vetores, respetivamente, de espaços vetoriais V e W sobre um corpo arbitrário $\mathbb{F}$, $v \otimes w$ denota o produto tensorial de v por w.
- Dados v e w vetores, respetivamente, de espaços vetoriais V e W sobre um corpo arbitrário $\mathbb{F}$, $v \wedge w$ denota o tensor antissimétrico de v por w.

Índice

Introdução

A Matemática com todo o seu esplendor característico vive das constantes interrogações de quem a opera. A busca por novos resultados, demonstrações e teorias prende milhares em torno deste enigmático mundo que se enuncia em frente dos nossos olhos. Porém, não estaremos a ser demasiado apologistas da inovação ao invés de prestar mais atenção ao que até agora foi alcançado?

Ora é precisamente sobre isso que incide esta dissertação. Graças ao trabalho dos Professores J. A. Dias Da Silva e Y. O. Hamidoune foram aplicados argumentos de Álgebra Linear e Multilinear no estudo de resultados provenientes da Teoria dos Números. Como consequência deste trabalho conseguiram responder afirmativamente à conjetura colocada por Erdős e Heilbronn que afirmava que o conjunto das somas de dois elementos distintos de um subconjunto finito, A, de $\mathbb{Z}_p$, com p primo, tem cardinalidade maior ou igual a $\min\{p, 2|A| - 3\}$.

O primeiro grande e importante resultado estudado dá-nos um critério de minimilidade para o grau do polinómio mínimo da soma de Kronecker de dois operadores lineares f e g em função dos graus dos seus polinómios mínimos,

$$\deg P_{(f\otimes I + I\otimes g)} \geqslant \min\{p, \deg P_f + \deg P_g - 1\}.$$

Este teorema irá permitir deduzir um resultado que, dados dois subconjuntos finitos e não vazios, A e B, de $\mathbb{Z}_p$, relaciona a cardinalidade do conjunto das somas, $|A + B|$, com as cardinalidades dos conjuntos A e B. Este teorema é conhecido como o Teorema de Cauchy-Davenport,

$$|A + B| \geqslant \min\{p, |A| + |B| - 1\}.$$

Dado um corpo $\mathbb{F}$ de característica p e V um espaço vetorial de dimensão finita, d, sobre $\mathbb{F}$, o segundo grande teorema abordado nesta dissertação permite estimar inferiormente a dimensão do subespaço cíclico para a *derivada* de um dado operador linear f,

$$\dim \varphi_{Df}(v \wedge f(v) \wedge \ldots \wedge f^{m-1}(v)) \geqslant \min\{p, (\dim \varphi_f(v) - m)m + 1\}.$$

Daqui inferir-se-á a conjetura de Erdős e Heilbronn. Referimo-nos a estes resultados como os mais

significativos uma vez que são fundamentais para a obtenção dos notáveis avanços alcançados. Esta importância é-lhes igualmente concedida visto que são aqueles onde se pode observar de forma mais marcante e preponderante a presença das referidas técnicas de Álgebra Linear e Multilinear.

A tese está estruturada da seguinte forma:

No Capítulo 1 faz-se um enquadramento algébrico, onde se podem encontrar os resultados necessários para a fundamentação de toda a teoria abordada nos restantes capítulos. Estes assentam em conteúdos de Álgebra Linear bem como de Álgebra Multilinear tais como: Produtos Tensoriais, Aplicações Multilineares Antissimétricas, Produto de Kronecker , entre outros.

No Capítulo 2 estuda-se a primeira aplicação de Álgebra Multilinear em Teoria de Números, baseando-nos no artigo de J. A. Dias Da Silva e Y. O. Hamidoune [6]. Os autores estabeleceram um critério de minimalidade que apresenta uma nova demonstração para o famoso Teorema de Cauchy-Davenport. De forma a compreender plenamente a dimensão de tal trabalho são também expostas as demonstrações originais, a de Cauchy e a de Davenport. O salto cronológico das primeiras demonstrações para uma das mais recentes e inovadoras é propositado e pretende salientar de forma ainda mais evidente a simplificação inerente a estas técnicas.

No Capítulo 3 aprofunda-se o estudo da aplicação de técnicas de Álgebra Multilinear em resultados de Teoria dos Números culminando na exposição de um teorema, a partir do qual decorre a demonstração da conjetura de Erdős e Heilbronn. Tal como é devido, este excelente trabalho teve repercussões ainda mais recentes sendo uma das quais o Método Polinomial — também analisado neste último capítulo.

Por fim, é ainda de referir que em 2005 Terence Tao apresentou, com análise harmónica, uma nova prova do Teorema de Cauchy-Davenport ao aplicar uma nova forma do princípio da incerteza à transformada de Fourier, [17]. Posteriormente este trabalho sofreu alterações por Song Guo e Zhi-Wei Sun, [9], que conduziram à demonstração de uma extensão da conjetura de Erdős e Heilbronn.

1. Álgebra Linear e Multilinear

Esta dissertação irá incidir, tal como foi explicitado na Introdução, em técnicas de Álgebra Multilinear e, ao longo desta secção, far-se-á uma revisão dos conceitos e resultados chave para a compreensão da mesma. Para tal ser exequível é necessário também uma breve abordagem de alguns dos conceitos principais de Álgebra Linear que farão parte dos desenvolvimentos desta dissertação.

Estes e mais resultados podem ser encontrados em *"Handbook of Linear Algebra*, [11], *"Multilinear Algebra"* de Russell Merris, [15] e *"Matrix Analysis"* de Rajendra Bhatia, [2].

1.1. Resultados Preliminares

Definição 1.1.1. *Dada uma matriz $A \in \mathbb{C}^{n\times n}$, o polinómio característico de A é o polinómio mónico de grau n, com coeficientes em $\mathbb{C}$ definido por:*

$$c_A(z) = \det(zI - A) = (z - \lambda_1)^{n_1} \dots (z - \lambda_k)^{n_k}, \text{ onde } n_1 + \dots + n_k = n,$$

onde n_i representa a multiplicidade algébrica do valor próprio λ_i.

Definição 1.1.2. *Dada uma matriz $A \in \mathbb{C}^{n\times n}$, um polinómio $p(z) = c_n z^n + \dots + c_1 z + c_0, c_n \neq 0 \in \mathbb{C}[z]$ diz-se um polinómio anulador da matriz A se $p(A) = c_n A^n + \dots + c_1 A + c_0 I_n$ representar a matriz nula.*

Definição 1.1.3. *Dada uma matriz $A \in \mathbb{C}^{n\times n}$, o polinómio mínimo de A, $p_A(z)$, define-se como o polinómio mónico de menor grau que anula a matriz A.*

Observação 1.1.1. *Note-se que quando uma aplicação é diagonalizável, as multiplicidades algébricas de todos os seus valores próprios coincidem com as multiplicidades geométricas. Nestas condições, o grau do polinómio mínimo coincide com o número de valores próprios distintos.*

Teorema 1.1.1. *Teorema de Cayley-Hamilton*
Dada uma matriz $A \in M_n$ e a matriz identidade de ordem n, I_n, o polinómio característico é anulador da matriz A, isto é,

$$c_A(A) = 0, \text{ com } c_A(\lambda) = \det(\lambda I_n - A).$$

Demonstração. Cf. [12], Capítulo 2, Teorema 2.4.2. □

Graças ao Teorema de Cayley-Hamilton é imediato que, em dimensão finita, o polinómio mínimo de uma matriz existe sempre.

Recorde-se ainda uma proposição bastante importante na medida em que permite estabelecer uma correspondência unívoca entre matrizes e operadores lineares. Isto será útil para podermos falar em polinómio mínimo de um operador apesar de este apenas ter sido definido para matrizes. A sua prova invocará um teorema central em Álgebra Linear, o Teorema da Extensão Linear.

Teorema 1.1.2. *Teorema da Extensão Linear*
Sejam V e W espaços vetoriais de dimensão finita sobre um mesmo corpo $\mathbb{F}$. Seja $(e_1, \ldots, e_n)$ uma base de V. Se $w_1, \ldots, w_n \in W$ então existe uma única aplicação linear $f : V \to W$ tal que

$$f(e_i) = w_i \ \forall i \in \{1, \ldots, n\}.$$

Demonstração. Seja v um elemento arbitrário de V. Como $(e_1, \ldots, e_n)$ é uma base de V, existem escalares $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ tais que

$$v = \sum_{i=1}^{n} \alpha_i e_i.$$

Estes escalares estão univocamente determinados. Define-se

$$f(v) = \alpha_1 f(e_1) + \ldots + \alpha_n f(e_n) := \alpha_1 w_1 + \ldots + \alpha_n w_n.$$

A unicidade decorre de cálculos elementares. □

Proposição 1.1.1. *Sejam V e W espaços vetoriais de dimensões finitas sobre um corpo $\mathbb{F}$, com $\dim V = n$ e $\dim W = m$. Sejam β e γ bases de V e W, respetivamente. Dada uma matriz $A \in M_{m\times n}$, existe uma única aplicação linear $f : V \to W$ tal que A é a matriz de f em relação às bases β e γ, isto é, $A = M(f; \beta, \gamma)$.*

Demonstração. Suponha-se que

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

Pelo Teorema da Extensão Linear, Teorema 1.1.2, existe uma e uma só aplicação $f : V \to W$ tal

que

$$\begin{aligned} f(e_1) &= a_{11}v_1 + \ldots + a_{m1}v_m \\ &\vdots \\ f(e_n) &= a_{1n}v_1 + \ldots + a_{mn}v_m, \end{aligned}$$

isto é, tal que $M(f;\beta,\gamma) = A$. □

Proposição 1.1.2. *Sejam E, V e W espaços vetoriais de dimensão finita sobre $\mathbb{F}$. Sejam β, γ, δ bases de E, V e W, respetivamente, e $f : E \to V$ e $g : V \to W$ aplicações lineares. Assim,*

$$M(g \circ f; \beta, \delta) = M(g; \gamma, \delta)M(f; \beta, \gamma).$$

Demonstração. Considerem-se as bases $\beta = \{e_1, \ldots, e_n\}$, $\gamma = \{v_1, \ldots, v_m\}$ e $\delta = \{w_1, \ldots, w_p\}$ e as matrizes das aplicações lineares

$$A = M(f;\beta,\gamma) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \text{ e } \; B = M(g;\gamma,\delta) = \begin{bmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{p1} & \cdots & b_{pm} \end{bmatrix}.$$

Como $g \circ f : E \to W$ é linear, tem-se que $M(g \circ f; \beta, \delta) \in M_{p\times n}$. Adicionalmente,

$$M(g \circ f; \beta, \delta)_{kj} = d_{kj},$$

onde os coeficientes d_{kj} são definidos por

$$(g \circ f)(e_j) = \sum_{k=1}^{p} d_{kj} w_k,$$

para $1 \leqslant j \leqslant n$. Por outro lado,

$$g(f(e_j)) = g\Big(\sum_{i=1}^{m} a_{ij}v_i\Big) = \sum_{i=1}^{m} a_{ij}g(v_i) = \sum_{i=1}^{m} a_{ij}\Big(\sum_{k=1}^{p} b_{ki}w_k\Big) = \sum_{k=1}^{p}\Big(\sum_{i=1}^{m} b_{ki}a_{ij}\Big)w_k,$$

que, para $1 \leqslant j \leqslant n$, é precisamente $(g \circ f)(e_j)$.

Visto que $\{w_1, \ldots, w_p\}$ é linearmente independente resulta que

$$d_{kj} = \sum_{i=1}^{m} b_{ki}a_{ij}, 1 \leqslant k \leqslant p, 1 \leqslant j \leqslant n.$$

Por fim, $\sum_{i=1}^{m} b_{ki}a_{ij}$ é a entrada (k, j) do produto das matrizes B e A, $BA = M(g;\gamma,\delta)M(f;\beta,\gamma)$.

□

Proposição 1.1.3. *Sejam V e W espaços vetoriais de dimensão finita sobre $\mathbb{F}$. Sejam β_1, β_2 bases de V e γ_1, γ_2 bases de W. Se $f : V \to W$ é uma aplicação linear, então*

$$M(f; \beta_2, \gamma_2) = M(id_W; \gamma_1, \gamma_2)M(f; \beta_1, \gamma_1)M(id_V, \beta_2, \beta_1).$$

Demonstração. Da proposição anterior, Proposição 1.1.2, decorre que

$$M(f; \beta_2, \gamma_1) = M(f; \beta_1, \gamma_1)M(f; \mathrm{id}_V, \beta_2, \beta_1)$$

e, ainda que

$$M(f; \beta_2, \gamma_2) = M(\mathrm{id}_W; \gamma_1, \gamma_2)M(f; \beta_2, \gamma_1),$$

o que conclui a demonstração. □

A proposição seguinte é bastante significativa uma vez que, dela deriva a possibilidade de definição do polinómio característico de um operador linear f num espaço vetorial V, $c_f(t)$, como sendo $c_A(t)$ onde A é a matriz de f em relação a uma base arbitrária de V.

Proposição 1.1.4. *Sejam V um espaço vetorial de dimensão finita sobre um corpo $\mathbb{F}$ e β, γ bases de V sobre $\mathbb{F}$. Seja $f : V \to V$ uma aplicação linear. Considerem-se as seguintes matrizes: $A = M(f; \beta, \beta)$, $B = M(f; \gamma, \gamma)$ e $P = M(id_V; \beta, \gamma)$. Assim,*

$$A = P^{-1}BP,$$

isto é, matrizes do mesmo operador linear a respeito de bases eventualmente distintas são semelhantes.

Para além disso, matrizes semelhantes têm o mesmo polinómio característico.

Demonstração. Considerando $V = W$, $\beta_2 = \gamma_2 = \beta$ e $\beta_1 = \gamma_1 = \gamma$ na Proposição 1.1.3, obtém-se que

$$A = M(f; \beta, \beta) = M(\mathrm{id}_V; \gamma, \beta)M(f; \gamma, \gamma)M(\mathrm{id}_V; \beta, \gamma) = P^{-1}BP.$$

Adicionalmente, matrizes semelhantes têm o mesmo polinómio característico visto que

$$\begin{aligned} c_A(\lambda) &= \det(\lambda I - A) = \det(\lambda P^{-1}P - P^{-1}BP) \\ &= \det(P^{-1}(\lambda I - B)P) = \det P^{-1} \det(\lambda I - B) \det P \\ &= \det P^{-1}P \det(\lambda I - B) = c_B(\lambda) \end{aligned}$$

□

A definição de polinómio característico de f é assim independente da base escolhida, visto que $c_A(t)$=$c_{A'}(t)$, onde A' é a matriz de f em relação a uma base distinta da base utilizada para construir A.

Resumindo, a proposição anterior permite definir o polinómio característico de uma aplicação linear como sendo o polinómio característico de qualquer matriz que a represente e, para além disso, permite falar indistintamente em valores próprios de uma aplicação linear ou de uma matriz.

Avançando de seguida para uma exposição mais focada nos conceitos de Álgebra Multilinear, é de reforçar que esta é uma área da Matemática que desenvolve os resultados de Álgebra Linear baseando-se no conceito de tensor. Assim, é apenas natural que os resultados expostos sejam similares e fundamentados nos anteriores.

No que se segue, admitam-se que os espaços vetoriais sobre os quais se está a trabalhar são de dimensão finita e que os corpos têm caraterística zero, excetuando indicação em contrário.

Definição 1.1.4. *Espaço dual de um espaço vetorial e conjunto dual de uma base.*
O espaço dual de um determinado espaço vetorial V sobre um corpo $\mathbb{F}$, denotado por V^*, é o espaço vetorial das aplicações lineares de V em $\mathbb{F}$. Supondo que o espaço vetorial V admite uma base $\{e_i\}_{i\in I}$, define-se o conjunto dual da base $\{e_i\}_{i\in I}$ como $\{e^{*j}\}_{j\in I}$, onde

$$e^{*j} : V \to \mathbb{F} \text{ é a única aplicação linear que satisfaz } e^{*j}(e_i) = \delta_{i,j}.$$

No caso linear conhece-se o Teorema da Extensão Linear, um resultado bastante importante uma vez que afirma que uma aplicação linear fica completamente definida pelos valores que toma nos elementos da base do espaço de partida.

Proposição 1.1.5. *Sejam V um espaço vetorial e V^* o seu espaço dual. Considerem-se $\{e_i\}_{i\in I}$ uma base de V e $\{e^{*j}\}_{j\in I}$ o conjunto dual associado.*
Então $\{e^{*j}\}_{j\in I}$ é linearmente independente em V^* e, no caso em que a dimensão de V é finita, tem-se que $\{e^{*j}\}_{j\in I}$ é uma base de V^*.

Demonstração. Seja J um subconjunto finito e arbitrário de I e sejam $\{c_i\}_{i\in J}$ coeficientes tais que

$$\sum_{i\in J} c_i e_i^* = 0.$$

Assim, para cada $v \in V$, tem-se

$$\sum_{i\in J} c_i e_i^*(v) = 0.$$

O objetivo consiste em ver que os coeficientes c_i são todos nulos para qualquer i em J. Para tal,

tome-se $j \in J$ arbitrário e escolha-se um elemento de V, $v = e_j$. Deste modo,

$$0 = \sum_{i \in J} c_i e_i^*(e_j) = \sum_{i \in J} c_i \delta_{i,j} = c_j.$$

Logo $c_j = 0 \ \forall j \in J$ e, como este conjunto J foi tomado arbitrariamente, tem-se que $\{e^{*i}\}_{i \in I}$ é linearmente independente em V^*.

Resta agora ver que se a dimensão de V for finita tem-se de facto uma base de V^*.

Suponha-se que V tem dimensão n e sejam $(e_1, \ldots, e_n)$ uma base de V e $(e^{*1}, \ldots, e^{*n})$ o seu conjunto dual. Pelo que se viu anteriormente, sabe-se que $(e^{*1}, \ldots, e^{*n})$ é um conjunto linearmente independente. De modo a provar que é uma base é necessário verificar que todo o elemento de V^* se escreve à custa de $(e^{*1}, \ldots, e^{*n})$, isto é,

$$\forall f \in V^* \ \exists \ \alpha_1, \ldots, \alpha_n \in \mathbb{F} : f = \sum_{i=1}^{n} \alpha_i e^{*i}.$$

Ora, pelo Teorema 1.1.2, é suficiente provar que as duas aplicações lineares têm a mesma imagem nos elementos da base de V.

Considere-se, para cada $f \in V^*$, $\alpha_i = f(e_i)$.

Para cada e_j temos que:

$$\sum_{i=1}^{n} \alpha_i e^{*i}(e_j) = \sum_{i=1}^{n} f(e_i) e^{*i}(e_j) = \sum_{i=1}^{n} f(e_i) \delta_{i,j} = f(e_j),$$

o que conclui a demonstração. □

Definição 1.1.5. *Sejam $V_1, V_2, \ldots, V_m$ espaços vetoriais complexos, com $\dim V_i = n_i$ para $i = 1, \ldots, m$. Define-se o seu produto cartesiano como o conjunto*

$$V_1 \times V_2 \times \ldots \times V_m = \{(v_1, v_2, \ldots, v_m) : v_i \in V_i, 1 \leqslant i \leqslant m\}.$$

Com as operações usuais de soma, e multiplicação por escalares o produto cartesiano $V_1 \times V_2 \times \ldots \times V_m$ é ainda um espaço vetorial sobre $\mathbb{F}$ com dimensão finita, sendo que $\dim(V_1 \times V_2 \times \ldots \times V_m) = n_1 + \ldots + n_m$.

Definição 1.1.6. *Considerem-se espaços vetoriais $V_1, V_2, \ldots, V_m, W$ sobre $\mathbb{F}$. Uma aplicação $f : V_1 \times V_2 \times \ldots \times V_m \to W$ diz-se multilinear se for linear sobre cada componente de $V_1 \times V_2 \times \ldots \times V_m$, isto é, se para quaisquer $\gamma, \delta \in \mathbb{F}, 1 \leqslant i \leqslant m$ se tem:*

$$f(v_1, v_2, \ldots, \gamma v_i + \delta w_i, \ldots, v_m) = \gamma f(v_1, v_2, \ldots, v_i, \ldots, v_m) + \delta f(v_1, v_2, \ldots, w_i, \ldots, v_m).$$

O espaço vetorial das aplicações multilineares de $V_1 \times V_2 \times \ldots \times V_m$ em W representa-se por $L(V_1, V_2, \ldots, V_m; W)$.

Ora, o interessante neste momento seria encontrar um resultado análogo ao Teorema da Extensão Linear para as aplicações multilineares. Para tal, recorde-se primeiramente a noção de base do produto cartesiano $V_1 \times \ldots \times V_m$.

Proposição 1.1.6. *Sejam $V_1, \ldots, V_m$ espaços vetoriais de dimensão finita, $\dim V_i = n_i, i = 1, \ldots, m$. Dadas bases $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ para os espaços V_i, constrói-se uma base para $V_1 \times \ldots \times V_m$:*

$$\{(\vec{0}, \ldots, \vec{0}, e_{ij}, \vec{0}, \ldots, \vec{0}) : j = 1, \ldots, n_i,\ i = 1, \ldots, m\}.$$

Considere-se $\rho : V_1 \times \ldots \times V_m \to W$ uma aplicação multilinear. Naturalmente se compreende que no caso das aplicações multilineares não é suficiente conhecer o valor de ρ nos vetores da base, visto que

$$\rho(v_1, \ldots, v_{i-1}, 0, v_{i+1}, \ldots, v_m) = 0.$$

Pretende-se então encontrar outro conjunto que desempenhe o mesmo papel que os vetores da base no caso linear e este será precisamente o seguinte:

$$\xi = \{(e_{1j_1}, \ldots, e_{mj_m}) : 1 \leqslant j_i \leqslant n_i,\ i = 1, \ldots, m\}.$$

Teorema 1.1.3. *Teorema da Extensão Multilinear*
Sejam $V_1, \ldots, V_m$ espaços vetoriais de dimensão finita sobre $\mathbb{F}$ com $\dim V_i = n_i,\ i = 1, \ldots, m$. Seja ξ o conjunto acima definido. Considerem-se os vetores $\{w_1, \ldots, w_{n_1 \times \ldots \times n_m}\}$ de um dado espaço vetorial W sobre o mesmo corpo $\mathbb{F}$. Então existe uma única aplicação multilinear

$$\rho : V_1 \times \ldots \times V_m \to W \text{ tal que}$$

$$\rho(e_{1j_1}, \ldots, e_{mj_m}) = w_k, \text{ onde } k \in \{1, \ldots, n_1 \times \ldots \times n_m\}.$$

Demonstração. Suponha-se que $v_i = \sum_{k=1}^{n_i} a_{ik} e_{ij_k}$, onde $v_i \in V_i$, para $i = 1, \ldots, m$. De forma a obter o pretendido, constrói-se a aplicação multilinear ρ definida da seguinte forma:

$$\begin{aligned}
\rho(v_1, \ldots, v_m) &= \rho\left(\sum_{k=1}^{n_1} a_{1k} e_{1j_k}, \ldots, \sum_{k=1}^{n_m} a_{mk} e_{mj_k}\right) = \\
&= \sum_{k=1}^{n_1} a_{1k} \times \ldots \times \sum_{k=1}^{n_m} a_{mk} \rho(e_{1j_k}, \ldots, e_{mj_k}) \\
&= \sum_{k=1}^{n_1} a_{1k} \times \ldots \times \sum_{k=1}^{n_m} a_{mk} \times w_k, \text{ onde } k \in \{1, \ldots, n_1 \times \ldots \times n_m\}.
\end{aligned}$$

Por construção, ρ é multilinear e satisfaz as condições pretendidas. A unicidade decorre de cálculos que usam o mesmo tipo de argumentação utilizada na prova da existência. □

1.2. Produto Tensorial

Os produtos tensoriais, a partir dos quais se definem os tensores antissimétricos, baseiam-se numa propriedade que recordaremos de seguida.

Definição 1.2.1. *Seja $\mathbb{T}$ um espaço vetorial e ψ uma função multilinear $\psi : V_1 \times V_2 \times \ldots \times V_m \to \mathbb{T}$. O par ($\mathbb{T}$,$\psi$) satisfaz a Propriedade de Fatorização Universal se para todo o espaço vetorial W e para toda a função multilinear $f : V_1 \times V_2 \times \ldots \times V_m \to W$ existir uma função linear $h : \mathbb{T} \to W$ tal que $f = h \circ \psi$.*

Definição 1.2.2. *O alcance de uma função f é o subespaço vetorial gerado pelo seu conjunto imagem, isto é, $< Imf >$.*

Com isto, enuncia-se a definição de produto tensorial.

Definição 1.2.3. *Escreve-se $\mathbb{T} = V_1 \otimes V_2 \otimes \ldots \otimes V_m$ e referimo-nos a $\psi(v_1, v_2, \ldots, v_m)$ como um tensor decomponível, $v_1 \otimes \ldots \otimes v_m$, se $\mathbb{T}$ for o produto tensorial dos espaços $V_1, V_2, \ldots, V_m$, isto é, se o par ($\mathbb{T}$,ψ) satisfizer a Propriedade de Fatorização Universal e o alcance da função $\psi : V_1 \times V_2 \times \ldots \times V_m \to \mathbb{T}$ for precisamente o espaço vetorial $\mathbb{T}$.*

Um produto tensorial pode também ser definido para espaços de aplicações lineares.

Proposição 1.2.1. *Sejam V_i e W_i, $i = 1, 2$, espaços vetoriais sobre um corpo $\mathbb{F}$. Considere-se $f_i \in L(V_i; W_i), i = 1, 2$, aplicações lineares. Existe uma única aplicação linear*

$$T : V_1 \otimes V_2 \to W_1 \otimes W_2$$

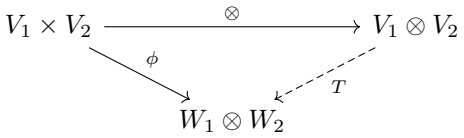
que verifica

$$T(v_1 \otimes v_2) = f_1(v_1) \otimes f_2(v_2).$$

Nestas condições denota-se T por $f_1 \otimes f_2$. Esta notação é utilizada uma vez que esta aplicação pode ser interpretada como o produto tensorial das aplicações f_1 e f_2 num modelo adequado de produto tensorial.

Demonstração. Seja $\phi : V_1 \times V_2 \to W_1 \otimes W_2$ tal que $\phi(v_1, v_2) = f_1(v_1) \otimes f_2(v_2)$. De acordo com a forma como foi definida, ϕ é uma aplicação multilinear. Assim, pela propriedade de Fatorização

Universal existe uma única aplicação linear, T,



tal que $T : V_1 \otimes V_2 \to W_1 \otimes W_2$, onde

$$T(v_1 \otimes v_2) = \phi(v_1, v_2) = f_1(v_1) \otimes f_2(v_2).$$

□

Este teorema pode ser generalizado para o produto tensorial de n aplicações lineares.

O resultado seguinte permite que, de agora em diante, se mencione um produto tensorial de quaisquer espaços vetoriais sem mais preocupações.

Teorema 1.2.1. *Dados $V_1, \ldots, V_m$ espaços vetoriais sobre um corpo $\mathbb{F}$ existe um produto tensorial desses espaços.*

Demonstração. Cf. [14], Capítulo 1.2, Teorema 2.3 □

Desta forma, para o estudo dos produtos tensoriais ficar completo, resta apenas descrever uma base de um produto tensorial, que se chamará base induzida.

Proposição 1.2.2. *Sejam $V_1, \ldots, V_m$ espaços vetoriais de dimensão finita sobre $\mathbb{F}$ cujas dimensões são, respetivamente, $n_1, \ldots, n_m$. Seja $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ uma base de V_i, $i = 1, \ldots, m$. O conjunto*

$$\mathbb{B} = \{e_{1\alpha(1)} \otimes \ldots \otimes e_{m\alpha(m)} : \alpha \in \Gamma(n_1, \ldots, n_m)\}$$

é uma base de $V_1 \otimes \ldots \otimes V_m$ e pode também ser representada por $e^{\otimes}_{\alpha}$.

Demonstração. Seja $v_1 \otimes \ldots \otimes v_m \in V_1 \otimes \ldots \otimes V_m$, com $v_i = \sum_{j=1}^{n_i} a_{ij} e_{ij}$, para $i = 1, \ldots, m$. Pretende-se ver que este tensor é uma combinação linear de elementos de $\mathbb{B}$.
Ora,

$$\begin{aligned} v_1 \otimes \ldots \otimes v_m &= \left(\sum_{j=1}^{n_1} a_{1j} e_{1j}\right) \otimes \ldots \otimes \left(\sum_{j=1}^{n_m} a_{mj} e_{mj}\right) \\ &= \sum_{\alpha \in \Gamma(n_1, \ldots, n_m)} \prod_{i=1}^{m} a_{i\alpha(i)} (e_{1\alpha(1)} \otimes \ldots \otimes e_{m\alpha(m)}). \end{aligned}$$

Uma vez que os tensores decomponíveis geram $V_1 \otimes \ldots \otimes V_m$ conclui-se que qualquer vetor de $V_1 \otimes \ldots \otimes V_m$ é combinação linear de vetores de $\mathbb{B}$. Assim $\mathbb{B}$ é gerador de $V_1 \otimes \ldots \otimes V_m$.

Verifique-se que $\mathbb{B}$ é um conjunto linearmente independente.
Para tal, suponha-se que existe $\gamma \in \Gamma(n_1, \ldots, n_m)$ tal que $(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)})$ é combinação linear de elementos de $\mathbb{B}$ excluindo o elemento construido a partir de γ. Pelo que foi dito anteriormente, ξ define completamente qualquer aplicação multilinear, logo existe uma única aplicação $\psi : V_1 \times \ldots \times V_m \to \mathbb{F}$ tal que

$$\psi(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) = 1 \text{ e}$$

$$\psi(e_{1\alpha(1)}, \ldots, e_{m\alpha(m)}) = 0 \ \forall \alpha \in \Gamma(n_1, \ldots, n_m) \setminus \{\gamma\}.$$

Pela propriedade de fatorização universal, existe uma única aplicação linear $h : V_1 \otimes \ldots \otimes V_m \to \mathbb{F}$ tal que $h(v_1 \otimes \ldots \otimes v_m) = \psi(v_1, \ldots, v_m)$. Assim,

$$h(e_\alpha^\otimes) = h(e_{1\alpha(1)} \otimes \ldots \otimes e_{m\alpha(m)}) = \psi(e_{1\alpha(1)}, \ldots, e_{m\alpha(m)}) = 0 \ \forall \alpha \in \Gamma(n_1, \ldots, n_m) \setminus \{\gamma\}.$$

Por outro lado,

$$h(e_\gamma^\otimes) = h(e_{1\gamma(1)} \otimes \ldots \otimes e_{m\gamma(m)}) = \psi(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) = 1.$$

Obtém-se assim uma contradição pois h é linear. Assim, $e_\gamma^\otimes$ não pode ser combinação linear de $e_\alpha^\otimes \ \forall \alpha \in \Gamma(n_1, \ldots, n_m) \setminus \{\gamma\}$. □

1.3. Produto e Soma de Kronecker

O produto de Kronecker é uma operação matricial entre matrizes de qualquer tamanho, cujo resultado é uma matriz que pode ser definida por blocos. Defina-se o caso mais simples onde operam apenas duas matrizes.

Definição 1.3.1. *Dadas*

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in M_{m\times n} \ e \ B \in M_{p\times q},$$

o seu produto de Kronecker, denotado por $A \otimes B$, é dado por

$$A \otimes B = \left[\begin{array}{c|c|c} a_{11}B & \cdots & a_{1n}B \\ \hline \vdots & \ddots & \vdots \\ \hline a_{m1}B & \cdots & a_{mn}B \end{array}\right] \in M_{mp\times nq}.$$

Outros resultados sobre o Produto de Kronecker podem ser encontrados em [10].

A soma de Kronecker de matrizes é uma operação entre matrizes quadradas cuja definição, escrita à custa do produto de Kronecker, conduz a uma matriz quadrada que também pode ser definida por blocos.

Definição 1.3.2. *Sejam $A \in M_{m\times m}$ e $B \in M_{n\times n}$. A sua soma de Kronecker, denotada por $A \oplus B$, é dada por*

$$A \oplus B = A \otimes I_n + I_m \otimes B.$$

1.4. Aplicações Multilineares Antissimétricas

Assim, encontram-se estabelecidas as condições para continuar o nosso estudo, focando-nos agora nos tensores antissimétricos. Estes, tais como os tensores simétricos, provêm do espaço $L^m(V;W)$ que representa o conjunto das aplicações multilineares cujo domínio é o m-ésimo produto cartesiano do espaço vetorial de dimensão finita V. Nesta secção considere-se, tal como já foi feito anteriormente, que o corpo sobre o qual se trabalha tem característica nula, exceptuando casos com indicação em contrário.

Definição 1.4.1. *Uma aplicação multilinear $\psi \in L^m(V;W)$ diz-se antissimétrica se para qualquer permutação $\sigma \in S_m$,*

$$\psi(v_{\sigma(1)}, \ldots, v_{\sigma(m)}) = sgn(\sigma)\psi(v_1, \ldots, v_m).$$

Analogamente à notação utilizada acima, $L^m_A(V;W)$ denota o subespaço de $L^m(V;W)$ cujos elementos são as aplicações multilineares antissimétricas.

Proposição 1.4.1. *O conjunto das aplicações multilineares antissimétricas de $L^m(V;U)$ é um espaço vetorial.*

Demonstração. Sejam $\rho, \psi \in L^m(V;U)$ aplicações multilineares antissimétricas, um escalar $\alpha \in \mathbb{F}$ e uma permutação $\sigma \in S_m$.

$$\begin{aligned}(\rho+\psi)(v_{\alpha(1)}, \ldots, v_{\alpha(m)}) &= \rho(v_{\alpha(1)}, \ldots, v_{\alpha(m)}) + \psi(v_{\alpha(1)}, \ldots, v_{\alpha(m)}) \\ &= \operatorname{sgn}(\sigma)\rho(v_1, \ldots, v_m) + \operatorname{sgn}(\sigma)\psi(v_1, \ldots, v_m) \\ &= \operatorname{sgn}(\sigma)(\rho+\psi)(v_1, \ldots, v_m).\end{aligned}$$

E,

$$\begin{aligned}(\alpha\rho)(v_{\alpha(1)}, \ldots, v_{\alpha(m)}) &= \alpha(\rho(v_{\alpha(1)}, \ldots, v_{\alpha(m)})) \\ &= \alpha(\operatorname{sgn}(\sigma)\rho(v_1, \ldots, v_m)) \\ &= \operatorname{sgn}(\sigma)(\alpha\rho)(v_1, \ldots, v_m).\end{aligned}$$

□

Observação 1.4.1. *Uma das primeiras ilações que se pode retirar em relação às aplicações antissimétricas é de que uma repetição de elementos em qualquer coordenada implica imediatamente a anulação desse vetor, isto é, tomando uma aplicação antissimétrica $\rho : V_1 \times \ldots \times V_m \to W$, $\rho(v_1, \ldots, v, v, \ldots, v_m) = 0$.*

Demonstração. Seja $\rho : V_1 \times \ldots \times V_m \to W$ uma aplicação antissimétrica e considere-se que há coordenadas repetidas $(v_1, v_2, ..., v_m)$ tais que $v_i = v_j$. Sem perda de generalidade, suponha-se

que $v_1 = v_2$. Assim,

$$\rho(v_2, v_1, \ldots, v_m) = \operatorname{sgn}(2,1)\rho(v_1, v_2, \ldots, v_m).$$

Deste modo

$$\rho(v_1, v_2, \ldots, v_m) = -\rho(v_1, v_2, \ldots, v_m)$$

donde se conclui que

$$\rho(v_1, v_2, \ldots, v_m) = 0,$$

uma vez que estamos sobre um corpo de característica diferente de 2. □

Para além disso e, como seria expectável, existe também um teorema da Extensão Multilinear Antissimétrica.

Teorema 1.4.1. *Sejam V e W espaços vetoriais sobre um corpo $\mathbb{F}$, $E = (e_1, \ldots, e_n)$ uma base de V e $(y_\alpha)_{\alpha \in Q_{m,n}}$ uma família de vetores de W. Então existe uma única aplicação multilinear antissimétrica $\rho : \times^m V \to W$ tal que*

$$\rho(e_{\alpha(1)}, \ldots, e_{\alpha(m)}) = y_\alpha, \text{ com } \alpha \in Q_{m,n}.$$

Antes de se expor a demonstração deste resultado, introduza-se uma função auxiliar que se prova ser multilinear antissimétrica.

Proposição 1.4.2. *Sejam V e U espaços vetoriais e seja $(e_1, \ldots, e_n)$ uma base de V. Para todo o $\alpha \in Q_{m,n}$ e $u \in U$, a aplicação $\zeta_{\alpha,u} : \times^m V \to U$ definida como:*

$$\zeta_{\alpha,u}(v_1, \ldots, v_m) = \left(\sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \prod_{i=1}^{m} a_{i,\alpha\sigma(i)} \right) u,$$

onde $a_{i,j} \in A$, $A \in \mathbb{C}^{m \times n}$ e $v_i = \sum_{j=1}^{n} a_{i,j} e_j$, é uma aplicação multilinear antissimétrica.

Demonstração.

$$\zeta_{\alpha,u}(v_{\rho(1)}, \ldots, v_{\rho(m)}) = \left(\sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \prod_{i=1}^{m} a_{\rho(i),\alpha\sigma\rho(i)} \right) u.$$

Tomando $\sigma\rho(i) = \eta(i)$, vem que

$$\zeta_{\alpha,u}(v_{\rho(1)}, \ldots, v_{\rho(m)}) = \left(\sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \prod_{i=1}^{m} a_{\sigma^{-1}\eta(i),\alpha\eta(i)} \right) u.$$

Designando $\sigma^{-1}\eta(i)$ por j,

$$\begin{aligned}\zeta_{\alpha,u}(v_{\rho(1)},\ldots,v_{\rho(m)}) &= \operatorname{sgn}(\rho)\Big(\sum_{\sigma\in S_m}\operatorname{sgn}(\sigma)\prod_{j=1}^{m}a_{j,\alpha\sigma(j)}\Big)u\\ &= \operatorname{sgn}(\rho)\zeta_{\alpha,u}(v_1,\ldots,v_m).\end{aligned}$$

□

Demonstração. (Teorema 1.4.1)

Seja

$$\rho=\sum_{\alpha\in Q_{m,n}}\zeta_{\alpha,y_\alpha},\ \text{onde}\ \zeta_{\alpha,y_\alpha}:\times^m V\to W.$$

A aplicação ρ é multilinear antissimétrica pois é a soma de aplicações multilineares antissimétricas e o conjunto $L_A^m(V;W)$ é um espaço vetorial.
Pretende-se mostrar que

$$\rho(e_{\beta(1)},\ldots,e_{\beta(m)})=y_\beta\ \forall\beta\in Q_{m,n}.$$

Ora,

$$\begin{aligned}\rho(e_{\beta(1)},\ldots,e_{\beta(m)}) &= \sum_{\alpha\in Q_{m,n}}\zeta_{\alpha,y_\alpha}(e_{\beta(1)},\ldots,e_{\beta(m)})\\ &= \sum_{\alpha\in Q_{m,n}}\Big(\sum_{\sigma\in S_m}\operatorname{sgn}(\sigma)\Big(\prod_{i=1}^{m}\delta_{\beta(i),\alpha\sigma(i)}\Big)\Big)y_\alpha\\ &= \sum_{\alpha\in Q_{m,n}}\Big(\sum_{\sigma\in S_m}\operatorname{sgn}(\sigma)\delta_{\beta,\alpha\sigma}\Big)y_\alpha=\Big(\sum_{\sigma\in S_m}\operatorname{sgn}(\sigma)\delta_{\beta,\beta\sigma}\Big)y_\beta\\ &= \Big(\sum_{\sigma\in H_\beta}\operatorname{sgn}(\sigma)\delta_{\beta,\beta\sigma}\Big)y_\beta=1\times y_\beta=y_\beta.\end{aligned}$$

Visto que

$$\delta_{\beta(i),\alpha\sigma(i)}=\begin{cases}1\text{ se }\beta=\alpha\sigma\iff\alpha=\beta,\\ 0\text{ caso contrário}\end{cases}$$

justificam-se assim as terceira e quarta igualdades. Seguidamente, o delta de Kronecker resultante, $\delta_{\beta,\beta\sigma}$, só fará sentido para as permutações que deixarem β invariante, isto é, para as que pertençam ao seu grupo estabilizador e por essa razão a soma passa a ser indexada apenas neste grupo, H_β. No entanto, a nossa permutação β, pertencente a $Q_{m,n}$, é estritamente crescente e portanto o único elemento que compõe o seu estabilizador será a permutação identidade. Assim visto que o sinal da permutação identidade é 1 obtém-se a penúltima igualdade.

Prove-se de seguida a unicidade.

Seja ψ uma aplicação multilinear antissimétrica tal que

$$\psi(e_\alpha) = y_\alpha \ \forall \alpha \in Q_{m,n}.$$

Sabe-se que para qualquer $\beta \in \Gamma(m,n)$ existem $\alpha \in G_{m,n}$ e $\sigma \in S_m$ tais que $\beta = \alpha\sigma$.

Note-se que se $\alpha \in G_{m,n} \setminus Q_{m,n}$, tanto o valor de $\psi(e_\alpha)$ como o de y_α será nulo, uma vez que nestas condições existiriam vetores iguais no argumento da função.

Logo,

$$\psi(e_{\beta(1)}, \ldots, e_{\beta(m)}) = \psi(e_{\alpha\sigma(1)}, \ldots, e_{\alpha\sigma(m)}).$$

Visto que ψ é uma função multilinear antissimétrica, da expressão anterior vem que

$$\psi(e_{\beta(1)}, \ldots, e_{\beta(m)}) = \operatorname{sgn}(\sigma)\psi(e_{\alpha(1)}, \ldots, e_{\alpha(m)}).$$

Pelo Teorema da Extensão Multilinear, 1.1.3, sabe-se que

$$\psi(e_\beta) = \rho(e_\beta) \ \forall \beta \in \Gamma(m,n).$$

Portanto

$$\psi(e_{\beta(1)}, \ldots, e_{\beta(m)}) = \operatorname{sgn}(\sigma)\rho(e_{\alpha(1)}, \ldots, e_{\alpha(m)}).$$

Contudo, ρ é também uma aplicação multilinear antissimétrica e, consequentemente, tem-se

$$\psi(e_{\beta(1)}, \ldots, e_{\beta(m)}) = \operatorname{sgn}(\sigma)\rho(e_{\alpha(1)}, \ldots, e_{\alpha(m)}) = \rho(e_{\alpha\sigma(1)}, \ldots, e_{\alpha\sigma(m)}) = \rho(e_{\beta(1)}, \ldots, e_{\beta(m)}).$$

□

Verificam-se agora as condições necessárias para definir o espaço de Grassmann de grau m, também conhecido por Espaço dos Tensores Antissimétricos de grau m:

Definição 1.4.2. *Sejam V e W espaços vetoriais sobre $\mathbb{F}$ e ξ uma aplicação multilinear antissimétrica de $\times^m V$ em W. Dizemos que (ξ, W) é um espaço de tensores antissimétricos associado a V se o alcance do operador ξ for W e ainda se se verificar a Propriedade de Fatorização Universal.*

Como é natural esta propriedade está em tudo relacionada com a que já tinha sido introduzida anteriormente.

Observação 1.4.2. *Propriedade de fatorização universal para o espaço de Grassmann:*
Diz-se que (ξ, W) verifica a Propriedade de Fatorização Universal se para qualquer espaço vetorial U e qualquer aplicação multilinear antissimétrica $\psi : \times^m V \to U$ existe uma única aplicação linear $h : W \to U$ tal que $\psi = (h \circ \xi)$. Assim, podemos escrever o seguinte

$$\psi(v_1, \ldots, v_m) = h(v_1 \wedge \ldots \wedge v_m).$$

Os tensores $\xi(v_1,\ldots,v_m) = v_1 \wedge \ldots \wedge v_m$ *dizem-se tensores decomponíveis antissimétricos e denotamos* $U = \underbrace{V \wedge \ldots \wedge V}_{m \text{ vezes}} := \bigwedge^m V$.

Vamos, no que se segue, usar como modelo para o espaço $\bigwedge^m V$ o seguinte:

$$v_1 \wedge \ldots \wedge v_m = \frac{1}{m!} \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma)(v_{\sigma^{-1}(1)} \otimes \ldots \otimes v_{\sigma^{-1}(m)}).$$

Proposição 1.4.3. *Sejam* V *um espaço vetorial de dimensão finita,* n, $v_1,\ldots,v_m$ *vetores de* V *e* $A = (a_{ij}) \in \mathbb{F}^{m\times n}$ *a matriz tal que* $v_i = \sum_{j=1}^n a_{ij}e_j$, *com* $i = 1,\ldots,m$. *Então*

$$v_1 \wedge \ldots \wedge v_m = \sum_{\alpha \in Q_{m,n}} \det A[id|\alpha] e_\alpha^\wedge.$$

De forma a ser possível demonstrar a proposição anterior, vejamos primeiro as duas proposições seguintes e as suas demonstrações.

Proposição 1.4.4. *Sejam* V *um espaço vetorial de dimensão finita,* $f_1,\ldots,f_m$ *operadores lineares em* V, *isto é,* $f_1,\ldots,f_m \in V^*$ *e* $v_1,\ldots,v_m$ *vetores de* V. *Então*

$$m!(f_1 \otimes \ldots \otimes f_m)(v_1 \wedge \ldots \wedge v_m) = \det \begin{bmatrix} f_1(v_1) & \cdots & f_m(v_1) \\ \vdots & \ddots & \vdots \\ f_1(v_m) & \cdots & f_m(v_m) \end{bmatrix}.$$

Demonstração. Sejam $f_1,\ldots,f_m \in V^*$ e $v_1,\ldots,v_m \in V$. Pelo modelo considerado, temos que

$$m!(f_1 \otimes \ldots \otimes f_m)(v_1 \wedge \ldots \wedge v_m) = (f_1 \otimes \ldots \otimes f_m)\left(\frac{1}{m!} \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma)(v_{\sigma^{-1}(1)} \otimes \ldots \otimes v_{\sigma^{-1}(m)})\right).$$

Como $f_1 \otimes \ldots \otimes f_m$ é linear vem que

$$\begin{aligned}
(f_1 \otimes \ldots \otimes f_m)(v_1 \wedge \ldots \wedge v_m) &= \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma)(f_1 \otimes \ldots \otimes f_m)(v_{\sigma^{-1}(1)} \otimes \ldots \otimes v_{\sigma^{-1}(m)}) \\
&= \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \times \left(f_1(v_{\sigma^{-1}(1)}) \times \ldots \times f_m(v_{\sigma^{-1}(m)})\right) \\
&= \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \prod_{i=1}^m f_i(v_{\sigma^{-1}(i)}) \\
&= \det \begin{bmatrix} f_1(v_1) & \cdots & f_m(v_1) \\ \vdots & \ddots & \vdots \\ f_1(v_m) & \cdots & f_m(v_m) \end{bmatrix}.
\end{aligned}$$

□

Proposição 1.4.5. *Considerem-se V um espaço vetorial de dimensão finita, n, e $(e_1, \ldots, e_n)$ uma base de V. Tomando a base dual, $(e^{*1}, \ldots, e^{*n})$, e $\alpha, \beta \in Q_{m,n}$ tem-se que*

$$e_\alpha^{*\otimes}(e_\beta^\wedge) = (e^{*\alpha(1)} \otimes \ldots \otimes e^{*\alpha(m)})(e_{\beta(1)} \wedge \ldots \wedge e_{\beta(m)}) = \frac{1}{m!}\delta_{\alpha,\beta}.$$

Demonstração. Sejam $(e_1, \ldots, e_n)$ uma base de V, $(e^{*1}, \ldots, e^{*n})$ a sua base dual e $\alpha, \beta \in Q_{m,n}$. Ora, pelo modelo considerado para o espaço $\bigwedge^m V$, temos que

$$e_\beta^\wedge = e_{\beta(1)} \wedge \ldots \wedge e_{\beta(m)} = \frac{1}{m!} \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma)(e_{\beta(\sigma^{-1}(1))} \otimes \ldots \otimes e_{\beta(\sigma^{-1}(m))}).$$

Assim,

$$\begin{aligned}
e_\alpha^{*\otimes}(e_\beta^\wedge) &= \big(e^{*\alpha(1)} \otimes \ldots \otimes e^{*\alpha(m)}\big)\bigg(\frac{1}{m!} \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma)\big(e_{\beta(\sigma^{-1}(1))} \otimes \ldots \otimes e_{\beta(\sigma^{-1}(m))}\big)\bigg) \\
&= \frac{1}{m!} \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma)\bigg(\big(e^{*\alpha(1)} \otimes \ldots \otimes e^{*\alpha(m)}\big)\big(e_{\beta(\sigma^{-1}(1))} \otimes \ldots \otimes e_{\beta(\sigma^{-1}(m))}\big)\bigg) \\
&= \frac{1}{m!} \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma)\bigg(\prod_{i=1}^{m} e^{*\alpha(i)}\big(e_{\beta(\sigma^{-1}(i))}\big)\bigg) = \frac{1}{m!} \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma)\delta_{\alpha,\beta\sigma^{-1}} \\
&= \frac{1}{m!}\bigg(\sum_{\sigma \in S_m} \operatorname{sgn}(\sigma)\delta_{\alpha,\alpha\sigma^{-1}}\bigg)\delta_{\alpha,\beta} = \frac{1}{m!}\bigg(\sum_{\sigma \in H_\alpha} \operatorname{sgn}(\sigma)\delta_{\alpha,\alpha\sigma^{-1}}\bigg)\delta_{\alpha,\beta},
\end{aligned}$$

onde penúltima e última igualdades são justificadas pelo mesmo argumento que foi utilizado na demonstração do Teorema 1.4.1.

Uma vez que $\alpha \in Q_{m,n}$, o único elemento do seu estabilizador será a permutação identidade. Logo,

$$\begin{aligned}
e_\alpha^{*\otimes}(e_\beta^\wedge) &= \frac{1}{m!}\Big(\operatorname{sgn}(\mathrm{id})\delta_{\alpha,\alpha\mathrm{id}^{-1}}\Big)\delta_{\alpha,\beta} = \frac{1}{m!}\Big(1 \times \delta_{\alpha,\alpha\mathrm{id}}\Big)\delta_{\alpha,\beta} = \frac{1}{m!}\Big(\delta_{\alpha,\alpha}\Big)\delta_{\alpha,\beta} \\
&= \frac{1}{m!}\delta_{\alpha,\beta}
\end{aligned}$$

□

Estamos agora em condições de demonstrar a Proposição 1.4.3.

Demonstração. Demonstração da Proposição 1.4.3.

Sejam $(e_1, \ldots, e_n)$ uma base de V, considere-se $(e^{*1}, \ldots, e^{*n})$ a sua base dual, os vetores $v_1, \ldots, v_m$ de V e $\alpha \in Q_{m,n}$.

Ora o objetivo é encontrar precisamente a expressão que defina os coeficientes c_α da expressão

seguinte:

$$v_1 \wedge \ldots \wedge v_m = \sum_{\alpha \in Q_{m,n}} c_\alpha e_\alpha^{\wedge}.$$

Por um lado,

$$\begin{aligned} e_\alpha^{*\otimes}(v_1 \wedge \ldots \wedge v_m) &= (e^{*\alpha(1)} \otimes \ldots \otimes e^{*\alpha(m)})(v_1 \wedge \ldots \wedge v_m) \\ &= (e^{*\alpha(1)} \otimes \ldots \otimes e^{*\alpha(m)})\left(\sum_{\beta \in Q_{m,n}} c_\beta e_\beta^{\wedge} \right) = \sum_{\beta \in Q_{m,n}} c_\beta e_\alpha^{*\otimes}(e_\beta^{\wedge}). \end{aligned}$$

Assim, utilizando a Proposição 1.4.5, vem que:

$$e_\alpha^{*\otimes}(v_1 \wedge \ldots \wedge v_m) = \sum_{\beta \in Q_{m,n}} c_\beta \frac{1}{m!} \delta_{\alpha,\beta} = c_\alpha \frac{1}{m!}$$

Por outro lado, pela Proposição 1.4.4, temos que:

$$e_\alpha^{*\otimes}(v_1 \wedge \ldots \wedge v_m) = (e^{*\alpha(1)} \otimes \ldots \otimes e^{*\alpha(m)})(v_1 \wedge \ldots \wedge v_m) = \frac{1}{m!} \det \begin{bmatrix} e^{*\alpha(1)}(v_1) & \cdots & e^{*\alpha(m)}(v_1) \\ \vdots & \ddots & \vdots \\ e^{*\alpha(1)}(v_m) & \cdots & e^{*\alpha(m)}(v_m) \end{bmatrix}.$$

Recordando que $e^{*j}(e_i) = \delta_{i,j}$ para $j, i = 1, \ldots, m$, vem que para todo $i, k \in \{1, \ldots, m\}$,

$$e^{*\alpha(i)}(v_k) = e^{*\alpha(i)}\left(\sum_{j=1}^{n} a_{kj} e_j \right) = \sum_{j=1}^{n} a_{kj} e^{*\alpha(i)}(e_j) = \sum_{j=1}^{n} a_{kj} \delta_{j,\alpha(i)} = a_{k,\alpha(i)}.$$

Consequentemente,

$$\begin{aligned} e_\alpha^{*\otimes}(v_1 \wedge \ldots \wedge v_m) &= (e^{*\alpha(1)} \otimes \ldots \otimes e^{*\alpha(m)})(v_1 \wedge \ldots \wedge v_m) \\ &= \det \begin{bmatrix} a_{1\alpha(1)} & \cdots & a_{1\alpha(m)} \\ \vdots & \ddots & \vdots \\ a_{m\alpha(1)} & \cdots & a_{m\alpha(m)} \end{bmatrix} = \det A[\mathrm{id}|\alpha]. \end{aligned}$$

Assim, resulta que

$$c_\alpha \frac{1}{m!} = \frac{1}{m!} \det A[\mathrm{id}|\alpha],$$

isto é,

$$c_\alpha = \det A[\mathrm{id}|\alpha], \text{ tal como pretendido.}$$

□

Resta agora definir, de forma pouco surpreendente, uma base para o espaço de Grassmann.

Proposição 1.4.6. *Considerando ainda as condições acima e $E = (e_1, \ldots, e_n)$ uma base de V, o conjunto*

$$\mathbb{D} = \{e_{\alpha(1)} \wedge \ldots \wedge e_{\alpha(m)} : \alpha \in Q_{m,n}\}$$

constitui uma base de $\bigwedge^m V$.

Demonstração. Seja $v_1 \wedge \ldots \wedge v_m \in \bigwedge^m V$, com $v_i = \sum_{j=1}^n a_{ij}e_j$, para $i = 1, \ldots, m$. Pretende-se ver que este tensor é uma combinação linear de elementos de $\mathbb{D}$.
Ora, pela Proposição 1.4.3 tem-se que

$$\begin{aligned} v_1 \wedge \ldots \wedge v_m &= \left(\sum_{j=1}^n a_{1j}e_j\right) \wedge \ldots \wedge \left(\sum_{j=1}^n a_{mj}e_j\right) \\ &= \sum_{\alpha \in Q_{m,n}} \det A[\mathrm{id}|\alpha](e_{\alpha(1)} \wedge \ldots \wedge e_{\alpha(m)}). \end{aligned}$$

Uma vez que os tensores decomponíveis geram $\bigwedge^m V$ conclui-se que qualquer vetor de $\bigwedge^m V$ é combinação linear de vetores de $\mathbb{D}$. Assim $\mathbb{D}$ é gerador de $\bigwedge^m V$.
Verifique-se que $\mathbb{D}$ é um conjunto linearmente independente. Para tal, suponha-se que existe $\gamma \in Q_{m,n}$ tal que $(e_{\gamma(1)} \wedge \ldots \wedge e_{\gamma(m)})$ é combinação linear de elementos de $\mathbb{D}$ excluindo o elemento construído a partir de γ. Pelo que foi dito anteriormente, $(e_{\alpha(1)}, \ldots, e_{\alpha(m)})$, com $\alpha \in Q_{m,n}$, define completamente qualquer aplicação multilinear antissimétrica, logo existe uma única aplicação $\psi : \times^m V \to \mathbb{F}$ tal que

$$\psi(e_{\gamma(1)}, \ldots, e_{\gamma(m)}) = 1 \text{ e}$$

$$\psi(e_{\alpha(1)}, \ldots, e_{\alpha(m)}) = 0 \ \forall \alpha \in Q_{m,n} \setminus \{\gamma\}.$$

Pela propriedade de fatorização universal, existe uma única aplicação linear $h : \bigwedge^m V \to \mathbb{F}$ tal que $h(v_1 \wedge \ldots \wedge v_m) = \psi(v_1, \ldots, v_m)$. Assim,

$$h(e_\alpha^\wedge) = h(e_{\alpha(1)} \wedge \ldots \wedge e_{\alpha(m)}) = \psi(e_{\alpha(1)}, \ldots, e_{\alpha(m)}) = 0 \ \forall \alpha \in Q_{m,n} \setminus \{\gamma\}.$$

Por outro lado,

$$h(e_\gamma^\wedge) = h(e_{\gamma(1)} \wedge \ldots \wedge e_{\gamma(m)}) = \psi(e_{\gamma(1)}, \ldots, e_{\gamma(m)}) = 1.$$

Obtém-se assim uma contradição pois, como h é linear, $e_\gamma^\wedge$ não pode ser combinação linear de $e_\alpha^\wedge \ \forall \alpha \in Q_{m,n} \setminus \{\gamma\}$. □

Assim, quando V tem dimensão finita d, e $1 \leqslant m \leqslant d$, torna-se claro que

$$\dim\left(\bigwedge{}^m V\right) = \binom{d}{m}$$

e se $m > d$ então $\dim \bigwedge^m V = 0$.

Por fim, apresentam-se algumas proposições cuja utilização posterior permitirá conduzir à dedução dos teoremas principais abordados nesta tese:

Proposição 1.4.7. *Sejam $v_1,\ldots,v_m$ elementos de um espaço vetorial V, com $m \leqslant \dim V$. Então $v_1 \wedge \ldots \wedge v_m = 0$ se, e somente se, $(v_1,\ldots,v_m)$ for linearmente dependente.*

Demonstração. Suponha-se que $(v_1,\ldots,v_m)$ é um conjunto linearmente dependente.
Assim $\exists j \in \{1,\ldots,m\} : v_j = a_1v_1 + \ldots + a_{j-1}v_{j-1} + a_{j+1}v_{j+1} + \ldots + a_mv_m$ com $a_i \in \mathbb{R},\ i \in \{1,\ldots,m\} \setminus \{j\}$. Sem perda de generalidade suponha-se que o primeiro elemento é combinação linear dos restantes.
Logo $v_1 = a_2v_2 + \ldots + a_mv_m$. Consequentemente,

$$\begin{aligned} v_1 \wedge v_2 \wedge \ldots \wedge v_m &= (a_2v_2 + \ldots + a_mv_m) \wedge v_2 \wedge \ldots \wedge v_m \\ &= a_2v_2 \wedge v_2 \wedge \ldots \wedge v_m + \ldots + a_mv_m \wedge v_2 \wedge \ldots \wedge v_m \\ &= 0 + \ldots + 0 = 0 \end{aligned}$$

A última igualdade deve-se à Observação 1.4.1.

Prove-se agora a outra implicação.
Seja $(e_1,\ldots,e_n)$ uma base de V e seja $A = (a_{ij}) \in \mathbb{C}^{m\times n}$ a matriz tal que $v_i = \sum_{j=1}^n a_{ij}e_j$. Pela Proposição 1.4.6 tem-se que

$$v_1 \wedge \ldots \wedge v_m = \sum_{\alpha\in Q_{m,n}} \det A[\mathrm{id}|\alpha]e_\alpha^\wedge.$$

Como, $v_1 \wedge \ldots \wedge v_m = 0$ e $e_\alpha^\wedge \neq 0$ (pela implicação anterior, visto que $e_\alpha^\wedge$ é uma base), vem que

$$\det A[\mathrm{id}|\alpha] = 0 \qquad \forall \alpha \in Q_{m,n}.$$

Deste modo as linhas da matriz A são linearmente dependentes, isto é, os vetores $v_1,\ldots,v_m$ são linearmente dependentes. □

A última proposição mostra que, geometricamente, todo o tensor não nulo $v_1 \wedge \ldots \wedge v_m$, corresponde a um subespaço vetorial de dimensão m de V.

Tal como em [11], Capítulo 13.8., tem-se que

Proposição 1.4.8. *Seja $\theta \in L(V;V)$ um operador diagonalizável. Seja $\bigotimes^m \theta$ a aplicação linear $\theta \otimes \ldots \otimes \theta$ de $\bigotimes^m V$ para $\bigotimes^m V$. A restrição de $\bigotimes^m \theta$ a $\bigwedge^m V$ será denotada por $\bigwedge^m \theta$.*

Se o espectro de θ for $\sigma(\theta) = \{\lambda_1,\ldots,\lambda_n\}$, então

$$\sigma\Big(\bigwedge{}^m\theta\Big) = \left\{\prod_{i=1}^m \lambda_{\alpha(i)}\right\}_{\alpha\in Q_{m,n}}.$$

Demonstração. Uma vez que θ é diagonalizável, considere-se $(e_1, \ldots, e_n)$ uma base de V constituída por vetores próprios de θ, onde e_i é o vetor próprio associado ao valor próprio λ_i. Assim,

$$\begin{aligned}
&\left(\bigwedge^m \theta\right)(e_{\alpha(1)}, \ldots, e_{\alpha(m)}) = (\theta \wedge \ldots \wedge \theta)(e_{\alpha(1)}, \ldots, e_{\alpha(m)}) \\
&= \theta(e_{\alpha(1)}) \wedge \ldots \wedge \theta(e_{\alpha(m)}) = \lambda_{\alpha(1)} e_{\alpha(1)} \wedge \ldots \wedge \lambda_{\alpha(m)} e_{\alpha(m)} \\
&= \prod_{i=1}^{m} \lambda_{\alpha(m)} (e_{\alpha(1)} \wedge \ldots \wedge e_{\alpha(m)}).
\end{aligned} \tag{1.1}$$

□

O estudo do segundo grande teorema desta tese requer o conhecimento de mais alguns conceitos para além dos que já foram introduzidos.

Definição 1.4.3. *Dado um espaço vetorial V de dimensão finita sobre um corpo $\mathbb{F}$, seja $\bigwedge^m V$ o m-ésimo espaço de Grassmann de V. Um operador linear f em V induz um operador linear Df, a derivada de f em $\bigwedge^m V$, definida por:*

$$Df(v_1 \wedge \ldots \wedge v_m) = \sum_{i=1}^{m} v_1 \wedge \ldots \wedge v_{i-1} \wedge f(v_i) \wedge v_{i+1} \wedge \ldots \wedge v_m.$$

Uma vez que os teoremas centrais desta dissertação assentam em conjuntos somas é necessário introduzir a seguinte notação referente aos mesmos.

Definição 1.4.4. *Seja A um subconjunto de um corpo $\mathbb{F}$, denota-se por $\bigwedge^m A$ o conjunto das somas de todos os subconjuntos de A de cardinalidade m.*

Exemplo 1.4.1. *Sejam $A = \{0, 2, 3, 4\}$ um subconjunto de $\mathbb{Z}_5$ e $m = 3$. Os subconjuntos de A de cardinalidade m são os seguintes:*

$$A_1 = \{0, 2, 3\} \qquad A_2 = \{0, 2, 4\} \qquad A_3 = \{0, 3, 4\} \qquad A_4 = \{2, 3, 4\}.$$

Assim,

$$\bigwedge^3 A = \{0, 1, 2, 4\}.$$

1.5. Partições e Diagrama de Young

As partições terão, tal como se comprovará, um papel muito importante na prova da conjetura de Erdős e Heilbronn. Assim, recordam-se de seguida algumas noções relacionadas com as mesmas.

Definição 1.5.1. *Uma sucessão de inteiros positivos $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_t)$ diz-se uma partição se $0 \leqslant \lambda_1 \leqslant \ldots \leqslant \lambda_t$.*

Neste estudo foi decidido definir partições como conjuntos cujos elementos se encontram ordenados de forma crescente.

Definição 1.5.2. *O comprimento de uma partição é o número de termos não nulos e denota-se por* $l(\lambda)$.

Denota-se o conjunto de todas as partições de comprimento menor ou igual a s como $\mathscr{P}_s$.

Definição 1.5.3. *O grau de uma partição* $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_s) \in \mathscr{P}_s$ *é*

$$\deg(\lambda) = \sum_{i=1}^{s} \lambda_i.$$

O conjunto de partições de grau k e comprimento no máximo s denota-se por $\mathscr{P}_{k,s}$. Usa-se a notação de s-tuplo para os elementos de $\mathscr{P}_{k,s}$ e, se necessário, faz-se uma identificação de uma partição com um s-tuplo com um conjunto de zeros no início.

Exemplo 1.5.1. *Considere-se o conjunto das partições de grau 4 e comprimento no máximo 3.*

$$\mathscr{P}_{4,3} = \{(1,1,2), (0,2,2), (0,1,3), (0,0,4)\}.$$

Uma vez que o conjunto acima representado contempla todas as partições de 4 *com comprimento no máximo* 3*, a identificação com um número de zeros no início possibilita que também as partições de comprimento inferior a* 3 *apareçam denotadas por triplos.*

Definição 1.5.4. *Sejam* $\lambda \in \mathscr{P}_{k+1,m}$ *e* $\mu \in \mathscr{P}_{k,m}$*. Escreve-se* $\mu \to \lambda$ *se existe um índice* j *tal que para todo o* i, $\lambda_i = \mu_i + \delta_{i,j}$*. Nestas condições diremos que* λ *segue a partir de* μ.

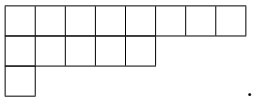
Analisando a definição anterior compreende-se que o facto de uma partição seguir a partir de outra equivale a dizer que estas apenas diferem em uma unidade em alguma componente.

As partições de grau k estão associadas de forma bijectiva com os caracteres irredutíveis de S_k (o grupo simétrico de grau k). Denotar-se-á por ξ_λ o carácter irredutível associado à partição λ.

Para além dos caracteres irredutíveis, a cada partição pode ser também associado um Diagrama de Young.

Definição 1.5.5. *Dada uma partição* $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_t)$*, o Diagrama de Young associado a* λ*, denotado por* $[\lambda]$*, consiste num diagrama composto por* $\deg(\lambda)$ *caixas dispostas em* t *linhas. Cada uma destas* t *linhas encontra-se alinhada a partir da mesma coluna, onde a linha* i *tem* λ_{t-i+1} *caixas,* $1 \leqslant i \leqslant t$.

Exemplo 1.5.2. *Como exemplo, representa-se o Diagrama de Young associado à partição* $\lambda = (1, 5, 8)$ *como o seguinte:*



.

Definição 1.5.6. *Considere-se uma partição* λ *de acordo com a definição acima. Na linha* i *e coluna* j *de* $[\lambda]$ *encontra-se a caixa* (i,j) *de* $[\lambda]$*. Referente à caixa* (i,j) *define-se* $H^{\lambda}_{i,j}$ *como o conjunto das caixas que se encontram na mesma linha para a direita da caixa* (i,j) *e para baixo na mesma coluna, incluindo também a respetiva caixa. O número de caixas de* $H^{\lambda}_{i,j}$ *representa-se por* $h^{\lambda}_{i,j}$*.*

Exemplo 1.5.3. *Considerando a partição tomada no Exemplo 1.5.2, tem-se que* $h^{\lambda}_{1,2} = 8$*.*

2. Teorema de Cauchy-Davenport

2.1. Álgebra Multilinear e o Teorema de Cauchy-Davenport

Com o intuito de estudar um dos grandes teoremas que foram mencionados anteriormente é necessário, antes de mais, estabelecer as condições sobre as quais a teoria será desenvolvida.

Ao longo deste capítulo considera-se V um espaço vetorial de dimensão finita sobre um corpo arbitrário $\mathbb{F}$. Denota-se por p a característica do corpo $\mathbb{F}$ caso este seja um corpo de característica não nula e seja $p = +\infty$ caso a característica de $\mathbb{F}$ seja nula.
Dado um operador linear f em V, o seu polinómio mínimo denota-se por P_f. Embora apenas se tenha definido, (em Definição 1.1.3), o polinómio mínimo de uma matriz, graças à observação que antecede a Proposição 1.1.4, é possível extender esta definição a operadores.

Introduz-se uma noção de graduação, baseada em subdivisões dos vectores que constituem um conjunto linearmente independente. De acordo com esta noção apresenta-se uma proposição que será fundamental nas demonstrações dos dois teoremas principais desta tese, uma vez que permite deduzir resultados de independência linear que não seriam tão diretos sem estes auxiliares.

Definição 2.1.1. *Seja A um conjunto de vetores linearmente independentes de V com uma decomposição $A = A_0 \dot\cup A_1 \dot\cup \ldots \dot\cup A_k$. Um vetor que seja combinação linear dos vetores de A tem grau i se for combinação linear dos vectores de $A_0 \dot\cup A_1 \dot\cup \ldots \dot\cup A_i$, com uma coordenada não nula num vetor de A_i.*

Proposição 2.1.1. *Seja B um conjunto de vetores linearmente independentes de V, com grau menor ou igual a i e v um vetor de grau $i+1$. Então $B \cup \{v\}$ é linearmente independente.*

Demonstração. Seja $B = \{v_1, \ldots, v_t\}$ um conjunto de vetores linearmente independentes de grau menor ou igual a i. Suponha-se, com vista a um absurdo, que $a \neq 0$, em

$$av + b_1 v_1 + \ldots + b_t v_t = 0. \tag{2.1}$$

Uma vez que v tem grau $i+1$ pode-se escrever como $v = v'_{i+1} + v'_i$, onde $v'_{i+1} \neq 0$ é combinação linear de vetores de A_{i+1} e v'_i é combinação linear de vetores de $A_0 \dot\cup A_1 \dot\cup \ldots \dot\cup A_i$. Portanto a

equação (2.1) pode ser rescrita como

$$av'_{i+1} + (av'_i + b_1v_1 + \ldots + b_tv_t) = 0.$$

Assim tem-se que

$$av'_{i+1} = -(av'_i + b_1v_1 + \ldots + b_tv_t),$$

o que é absurdo pois um vetor de grau $i+1$ não pode ser obtido como combinação linear de vetores de grau menor ou igual a i pela suposição da independência linear do conjunto A. □

Antes de podermos prosseguir para a apresentação do primeiro teorema central, defina-se o polinómio mínimo de um vetor em função de um operador.

Definição 2.1.2. *Sejam V um espaço vetorial de dimensão finita, n, sobre um corpo arbitrário $\mathbb{F}$ e f um operador linear em V.*
Visto que V tem dimensão finita, dado um vetor v, existe um inteiro m, com $0 \leqslant m \leqslant n$, o menor possível, tal que o vetor $f^m(v)$ é combinação linear de $v, f(v), \ldots f^{m-1}(v)$:

$$f^m(v) = \lambda_1 f^{m-1}(v) + ... + \lambda_m(v).$$

Define-se o polinómio mínimo de v em relação a f como o polinómio:

$$f^m(x) - \lambda_1 f^{m-1}(x) - ... - \lambda_m(x).$$

Visto isto, estamos agora em condições de apresentar o primeiro grande resultado estudado. O artigo no qual pode ser encontrado, [6], tem uma importância fulcral para a Matemática na medida em que foi de grande importância por iniciar uma linha de aplicação da Álgebra Multilinear à Teoria dos Números. Como será possível confirmar, esta aplicação não foi um caso único e a sua continuidade permitiu provar a tão famosa conjetura de Erdős e Heilbronn.

Antes da exposição do teorema referido, apresenta-se um lema e uma definição que serão bastante úteis no desenvolvimento do mesmo.

Lema 2.1.1. *Seja V um espaço vetorial de dimensão finita sobre um corpo arbitrário $\mathbb{F}$. Seja f um operador linear em V cujo grau do polinómio mínimo é k. Então existe em V um vetor, diga-se v, tal que o polinómio mínimo de v (em relação a f) é o polinómio mínimo de f.*

Demonstração. Cf. [8], Capítulo 7, Teorema 2. □

O Lema 2.1.1 permite então afirmar que, para um determinado $v \in V$, a família

$$(v, fv, f^2v, \ldots, f^{k-1}v)$$

é uma família de vetores linearmente independentes, caso contrário haveria um polinómio anulador para v de menor grau.

De modo a provar o critério de minimalidade estabelecido no enunciado do teorema seguinte será necessário introduzir um novo conceito, o peso.

Definição 2.1.3. *Sejam f e g operadores lineares sobre os espaços vetoriais de dimensão finita V e W, respetivamente. Define-se o peso de uma parcela do tipo $(f^i v \otimes g^j w)$ como o inteiro não negativo $i + j$ sempre que i e j não excedam o grau do polinómio mínimo dos respetivos operadores. Extrapolando a definição, diz-se que o tensor*

$$z \in \langle f^i v \otimes g^j w \mid 0 \leqslant i \leqslant k-1; 0 \leqslant j \leqslant r-1 \rangle$$

tem peso t quando

$$t = \max\{\textit{peso } f^i v \otimes g^j w \mid z \textit{ tem componente não nula em } f^i v \otimes g^j w\}.$$

Teorema 2.1.1. *Sejam V e W espaços vetoriais de dimensão finita sobre um corpo arbitrário $\mathbb{F}$ de característica p. Seja f um operador linear em V e seja g um operador linear em W. Então*

$$\deg P_{(f\otimes I + I\otimes g)} \geqslant \min\{p, \deg P_f + \deg P_g - 1\}.$$

Demonstração. Note-se, antes de mais, que o operador linear referente ao qual é estimado o grau do polinómio mínimo está bem definido graças à Proposição 1.2.1.

Suponha-se que $\deg P_f = k$ e $\deg P_g = r$.
Assim, existe $v \in V$ tal que $(v, fv, f^2v, \ldots, f^{k-1}v)$ é uma família de vetores linearmente independentes. Analogamente, existe $w \in W$ tal que $(w, gw, g^2w, \ldots, g^{r-1}w)$ é também uma família linearmente independente. A existência dos vetores v e w é garantida pelo Lema 2.1.1.

Considere-se o produto tensorial $V \otimes W$.

Já tinha sido observado que dadas bases de V e W, $\{e_1, \ldots, e_{\dim V}\}$ e $\{e'_1, \ldots, e'_{\dim W}\}$, respetivamente, obter-se-ia uma base de $V \otimes W$ construindo o seguinte conjunto:

$$\{e_i \otimes e'_j : 0 \leqslant i \leqslant \dim V; 0 \leqslant j \leqslant \dim W\}.$$

Tendo em conta que as famílias $(v, fv, \ldots, f^{k-1}v)$ e $(w, gw, g^2w, \ldots, g^{r-1}w)$ são linearmente independentes, é possível estendê-las a bases de V e W, respetivamente, e consequentemente, é possível concluir que $\mathscr{B} = (f^i v \otimes g^j w)$, com $0 \leqslant i \leqslant k-1$ e $0 \leqslant j \leqslant r-1$ é uma família linearmente independente, uma vez que é um subconjunto de uma base de $V \otimes W$.

Tendo presente o conceito de peso anteriormente enunciado, é possível demonstrar a seguinte proposição:

Proposição 2.1.2. *Se* $0 \leqslant m < \min\{p, k+r-1\}$*, então* $(f \otimes I + I \otimes g)^m v \otimes w$ *tem peso* m*.*

Demonstração. De forma a provar este resultado pretende-se encontrar uma forma mais prática de avaliar o peso da expressão anterior. Esta forma mais prática consistirá em escrever $(f \otimes I + I \otimes g)^m v \otimes w$ como combinação linear de elementos de $\mathscr{B}$ e, deste modo, a avaliação do peso total passará a ser feita parcela a parcela, de acordo com a definição dada.

Dado que $f \otimes I$ e $I \otimes g$ comutam, i.e, $(f \otimes I)(I \otimes g) = (f \otimes g) = (I \otimes g)(f \otimes I)$, tem-se que

$$\begin{aligned}(f \otimes I + I \otimes g)^m &= \sum_{q=0}^{m} \binom{m}{q} (f \otimes I)^q (I \otimes g)^{m-q} \\ &= \sum_{q=0}^{m} \binom{m}{q} (f^q \otimes I)(I \otimes g^{m-q}) = \sum_{q=0}^{m} \binom{m}{q} f^q \otimes g^{m-q}.\end{aligned}$$

Aplicando a igualdade de operadores aos vetores $v \otimes w$ tem-se que

$$(f \otimes I + I \otimes g)^m (v \otimes w) = \sum_{q=0}^{m} \binom{m}{q} f^q v \otimes g^{m-q} w. \tag{2.2}$$

Logo, com vista a provar que $(f \otimes I + I \otimes g)^m v \otimes w$ tem peso m, veja-se que

$$\sum_{q=0}^{m} \binom{m}{q} f^q v \otimes g^{m-q} w \text{ tem peso } m.$$

De modo a tornar a demonstração mais elucidativa, diga-se, de forma informal, que f tem expoente "grande" quando o expoente de f em $\binom{m}{q} f^q v \otimes g^{m-q} w$, q, é maior ou igual a k e que g tem expoente "grande" quando o seu expoente em $\binom{m}{q} f^q v \otimes g^{m-q} w$, $m-q$, é maior ou igual a r.

Divida-se esta análise em três casos:
(i) f com expoente "grande";
(ii) g com expoente "grande";
(iii) ambos f e g com expoentes "pequenos".

Antes de se proceder à análise de cada um dos casos acima apresentados note-se que estes são mutuamente exclusivos.
Ora se $q \geqslant k$, então

$$m - q \leqslant m - k < k + r - 1 - k = r - 1 < r, \text{ i.e, } m - q < r,$$

e caso $m - q \geqslant r$, tem-se

$$q \leqslant m - r < k + r - 1 - r < k - 1 < k, \text{ i.e, } q < k.$$

Logo, quando o expoente de f é "grande", o de g é "pequeno" e vice-versa. O caso (iii) é independente dos restantes visto que trata a hipótese onde ambos os expoentes são pequenos.

Caso (i): Se $q \geqslant k$, então

$$f^q v \in \langle v, fv, \dots, f^{q-1}v \rangle.$$

Portanto,

$$\sum_{q=0}^{m} \binom{m}{q} f^q v \otimes g^{m-q} w = \sum_{q=0}^{m} \binom{m}{q} (\alpha_1 v + \alpha_2 fv + \dots + \alpha_q f^{q-1} v) \otimes g^{m-q} w.$$

Assim, é facil inferir que, neste caso, o peso da soma será $q - 1 + m - q = m - 1 < m$.

Caso (ii): Efetuando um raciocínio análogo, se $m - q \geqslant r$, então

$$g^{m-q} w \in \langle w, gw, \dots, g^{m-q-1} w \rangle.$$

Deste modo,

$$\sum_{q=0}^{m} \binom{m}{q} f^q v \otimes g^{m-q} w$$

que pode ser rescrita como

$$\sum_{q=0}^{m} \binom{m}{q} f^q v \otimes (\gamma_1 w + \gamma_2 gw + \dots + \gamma_r g^{r-1} w)$$

terá também peso estritamente menor que m, visto que o seu peso é dado por
$q + r - 1 \leqslant q + m + q - 1 \leqslant m - 1 < m$.

Avaliados os pesos das parcelas em que f ou g têm expoentes "grandes", isto é, parcelas em que $q \geqslant k$ e em que $m - q \geqslant r$, respetivamente, concluiu-se que estas têm pesos estritamente menores que m.
Até agora, sabe-se portanto que o somatório em (2.2) tem dois tipos de parcelas com peso estritamente inferior a m e, de maneira a ser possível concluir o pretendido, resta então mostrar que existe um terceiro tipo de parcelas em que o peso das mesmas é exatamente m.

Ao estudar o primeiro caso, foram analisadas $m-k+1$ parcelas; no segundo, foram analisadas $m - r + 1$ parcelas, portanto das $m + 1$ parcelas do somatório em (2.2), restam analisar

$$(m + 1) - (m - k + 1) - (m - r + 1) \text{ parcelas.}$$

Ora

$$(m+1)-(m-k+1)-(m-r+1)=k-(m-r+1)$$

e

$$k-m+r-1>k-k-r+1+r-1=0.$$

Por conseguinte resta analisar um número positivo de parcelas. Isto mostra que

$$\sum_{q=0}^{m}\binom{m}{q}f^q v\otimes g^{m-q}w$$

tem outro conjunto de parcelas que não se encaixam em nenhum dos casos anteriores. Assim, sabe-se que as parcelas restantes são da forma $\binom{m}{q}f^q v\otimes g^{m-q}w$ com $0\leqslant q<k$ e $0\leqslant m-q<r$, ou seja, onde tanto f como g têm expoentes "pequenos".

Caso (iii) Uma vez garantida a existência de parcelas que se incluem neste caso, basta apenas determinar o seu peso e, para tal, é suficiente reparar que o mesmo será dado por $q+m-q=m$. Visto que estas parcelas não podem ser escritas à custa de nenhumas outras, o seu peso será calculado simplesmente através da soma dos expoentes.

De forma a dar por concluída esta parte da prova falta assegurar que pelo menos uma destas parcelas não se anula pois, caso contrário, as parcelas com peso m estariam associadas a coeficientes nulos e não seria possível demonstrar o pretendido. No entanto, uma subtil porém importante limitação do valor m permite responder a esta questão, que é precisamente o facto de que $m<p$. Se $m=p$, então $p\mid\binom{m}{q}$, logo haveria coeficientes nulos mod p; se $m\geq p$ não seria possível garantir que os coeficientes seriam não nulos. Portanto ao restringir m a um valor estritamente menor que p é garantido que todos os coeficientes são não nulos, concluindo assim a demonstração da Proposição 2.1.2. □

Por conseguinte, obtém-se então que $(f\otimes I+I\otimes g)^m v\otimes w$ tem peso m. Pretende-se de seguida ver que

$$((f\otimes I+I\otimes g)^i v\otimes w)_{0\leqslant i\leqslant min\{p,r+k-1\}-1}$$

é uma família linearmente independente. Para uma ideia mais clara desta família explicitem-se os primeiros três termos da mesma.

$$\begin{aligned}
&\big((f\otimes I+I\otimes g)^0 v\otimes w\big)=v\otimes w,\\
&\big((f\otimes I+I\otimes g)^1 v\otimes w\big)=fv\otimes w+v\otimes gw,\\
&((f\otimes I+I\otimes g)^2 v\otimes w)=f^2v\otimes w+fv\otimes gw+v\otimes g^2w.
\end{aligned}$$

Estes três exemplares são suficientes para poder depreender que para cada expoente i do operador $f\otimes I+I\otimes g$ se obtém uma combinação linear de vetores todos com o mesmo peso, i. Já vimos que

o conjunto destes vetores é linearmente independente. Assim, tomando como fator de graduação o conceito de peso, de acordo com a Proposição 2.1.1, decorre que a família

$$((f \otimes I + I \otimes g)^i v \otimes w)_{0 \leqslant i \leqslant \min\{p, k+r-1\}-1}$$

é de facto linearmente independente. Consequentemente,

$$\deg P_{f \otimes I + I \otimes g} \geqslant \min\{p, \deg P_f + \deg P_g - 1\}.$$

□

Tal como foi referido anteriormente, como consequência quase imediata, do Teorema 2.1.1, obtém-se uma demonstração alternativa para o Teorema de Cauchy-Davenport.

Propõe-se, tal como no teorema anterior, alcançar um critério de minimalidade para a cardinalidade de um determinado conjunto. Neste caso, o que se quer minorar é a cardinalidade do conjunto das somas de subconjuntos de $\mathbb{Z}_p$, sendo p um número primo.

Definição 2.1.4. *Seja p um número primo. Dados A e B subconjuntos não vazios de $\mathbb{Z}_p$ define-se o conjunto das somas de A e B como:*

$$|A + B| = \{a_i + b_j : a_i \in A \text{ e } b_j \in B\}.$$

Corolário 2.1.1 (Teorema de Cauchy-Davenport)**.** *Sejam A e B dois subconjuntos não vazios de $\mathbb{Z}_p$. Então*

$$|A + B| \geqslant \min\{p, |A| + |B| - 1\}.$$

Tendo em consideração o resultado anterior, a única subtileza que será necessária demonstrar será precisamente a prova de que o conjunto $A + B$ tem exatamente a mesma cardinalidade que o grau de um determinado polinómio mínimo. Esta dificuldade vai ser facilmente ultrapassada recorrendo novamente a instrumentos da Álgebra Multilinear. Se se provar que os elementos de $A + B$ se encontram na diagonal principal de uma matriz diagonal então o trabalho ficará concluído. Desta forma, recorra-se à soma de Kronecker de duas matrizes para tentar inferir o pretendido.

Demonstração. Sejam A e B dois subconjuntos não vazios de $\mathbb{Z}_p$. Suponha-se que $|A| = k$ e $|B| = r$, isto é, $A = \{a_1, \dots, a_k\}$ e $B = \{b_1 \dots, b_r\}$, onde os elementos de cada conjunto são distintos entre si.

Com vista a atingir o objetivo proposto, definem-se as seguintes matrizes $M = \operatorname{diag}(a_1, \dots, a_k)$ e $N = \operatorname{diag}(b_1, \dots, b_r)$. Uma vez que M e N são matrizes diagonais cujos elementos da diagonal são distintos, tem-se que o grau do respetivo polinómio mínimo é igual à cardinalidade de A e B, respetivamente, i.e., $\deg P_M = k$ e $\deg P_N = r$, de acordo com a Observação 1.1.1. Esta característica fundamenta naturalmente a escolha de matrizes diagonais.

Consequentemente, $M \otimes I_r + I_k \otimes N$ é uma matriz diagonal, com o seguinte aspeto $M \otimes I_r + I_k \otimes N =$

$$= \left[\begin{array}{c|c|c} a_1 & \mathbf{0} & \mathbf{0} \\ \ddots & & \\ a_1 & & \\ \hline \mathbf{0} & \ddots & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & a_k \\ & & \ddots \\ & & a_k \end{array}\right] + \left[\begin{array}{c|c|c} b_1 & \mathbf{0} & \mathbf{0} \\ \ddots & & \\ b_r & & \\ \hline \mathbf{0} & \ddots & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & b_1 \\ & & \ddots \\ & & b_r \end{array}\right]$$

$$= \left[\begin{array}{c|c|c} a_1 + b_1 & \mathbf{0} & \mathbf{0} \\ \ddots & & \\ a_1 + b_r & & \\ \hline \mathbf{0} & \ddots & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & a_k + b_1 \\ & & \ddots \\ & & a_k + b_r \end{array}\right].$$

Assim, tal como desejado, os elementos da diagonal principal são precisamente os elementos de $|A + B|$ e, portanto,

$$\deg P_{M\otimes I + I \otimes N} = |A + B|.$$

Aplicando agora uma versão matricial do teorema anterior, conclui-se que

$$|A + B| = \deg P_{M\otimes I+I\otimes N} \geqslant \min\{p, \deg P_M + \deg P_N - 1\} = \min\{p, |A| + |B| - 1\}.$$

□

É assim indubitável que, após a demonstração do Teorema 2.1.1, a demonstração do Teorema de Cauchy-Davenport se torna quase imediata. No entanto, para se compreender completamente o longo trabalho que a introdução de ferramentas de Álgebra Multilinear possibilitou contornar, é assim nosso dever compreender como foram as primeiras demonstrações deste teorema. As próximas secções incidem justamente sobre essas provas.

2.2. Prova de Cauchy do Teorema de Cauchy-Davenport

O teorema de Cauchy-Davenport, tal como o nome indica, recebeu este nome devido ao trabalho dos matemáticos Augustin-Louis Cauchy e Harold Davenport. No entanto, um facto curioso é que as datas das demonstrações têm mais de um século de diferença. Esta diferença cronológica deve-se ao facto de Davenport não ter tido conhecimento do trabalho desenvolvido por Cauchy aquando a sua publicação, dando conhecimento do mesmo no artigo [5]. Assim, seguindo a ordem natural, apresenta-se primeiramente a demonstração de Cauchy e posteriormente a de Davenport.

A primeira demonstração, feita por Cauchy, aparece em *"Recherches sur les nombres"*, [3] que data de 1813. Apesar de nesta altura já ter sido divulgada a linguagem modular, introduzida por Gauss em 1801, nesta demonstração os conceitos de aritmética modular ainda foram utilizados de uma forma tradicional e sem qualquer referência a *módulo p* ou a congruências propriamente ditas.

Na sua versão original o autor refere-se a um número p primo "escolhido" para divisor e, de acordo com esse divisor escolhido, dois números podem ter a mesma forma ou formas diferentes, referindo-se, respetivamente, a números congruentes e números não congruentes módulo p. Vamos então, à luz das noções da aritmética modular atuais, reproduzir as ideias latentes na demonstração original. Ao enunciar o teorema iremos utilizar a notação moderna.

Teorema 2.2.1. *Seja p um número primo e sejam A e B subconjuntos de $\mathbb{Z}_p$ com, respetivamente, $\alpha+1$ e $\beta+1$ elementos distintos tais que $\alpha+\beta+1 \leqslant p$. Então o conjunto $A+B$ terá, pelo menos, $\alpha+\beta+1$ elementos distintos de $\mathbb{Z}_p$.*

Antes de mais enunciem-se três resultados que serão auxiliares para a prova do Teorema 2.2.1.

Lema 2.2.1. *Considerem-se $\mathbb{Z}_p$, com p primo, e m e k números inteiros. O inteiro k é tal que $k \not\equiv 0 (mod\ p)$.*

Então $\mathbb{Z}_p = \{m+uk : 1 \leqslant u \leqslant p-1\}$.

O lema anterior revela que os elementos $m+uk$, com $u \in \{1,\ldots,p-1\}$, percorrem todas as classes módulo p.

Demonstração. Com vista a um absurdo, suponha-se que $m+rk \equiv m+sk$ para alguns inteiros $r,s \in \{1,\ldots,p-1\}$. Nestas condições, obter-se-ia que

$$rk \equiv sk \pmod{p} \iff k(r-s) \equiv 0 \pmod{p}.$$

Ora, visto que p é primo, tem-se que

$$p \mid k \vee p \mid r-s.$$

Como, $p \nmid k$ apenas poderia resultar que $p \mid r - s$ mas, como $|r - s| < p$ isto é igualmente impossível.

Assim, $m + uk$, com $u \in \{1, \ldots, p-1\}$, representa todas as classes de $\mathbb{Z}_p$. □

Teorema 2.2.2. *Sejam p um número primo e A um subconjunto não vazio de $\mathbb{Z}_p$ com $\alpha + 1$ elementos. Suponha-se que $\alpha + 2 \leqslant p$ e considere-se $k \not\equiv 0 \pmod{p}$.*
Tem-se que

$$|A \cup (A + k)| \geqslant \alpha + 2.$$

Demonstração. Suponha-se, com vista a um absurdo, que

$$|A \cup (A + k)| \leqslant \alpha + 1.$$

Para se obter um absurdo, será suficiente considerar que $A \cup (A + k)$ tem precisamente $\alpha + 1$ elementos, uma vez que $|A| = \alpha + 1$.
No entanto, de acordo com essa suposição, ter-se-á que $A = A + k$.
Assim, considere-se a_r um elemento arbitrário de A. Logo, $a_r + k \in A + k$.
Dado que o último conjunto coincide com A, existe um índice, distinto de r, diga-se s, tal que $a_r + k = a_s \in A$. Por sua vez, $a_r + 2k = a_s + k = a_t$ que ainda pertence a A. Deste modo, os seguintes elementos

$$a_r, a_r + k, a_r + 2k, a_r + 3k, \ldots, a_r + (p-1)k,$$

pertencem a A. De acordo com o Lema 2.2.1, tem-se ainda que estes números são distintos. Consequentemente, $|A| = p$, o que é um absurdo visto que $|A| = \alpha + 1 < p$. □

Teorema 2.2.3. *Seja p um número primo. Considere-se A um subconjunto de $\mathbb{Z}_p$ com $\alpha + 1$ elementos tal que $\alpha + 2 \leqslant p$. Sejam $b_0, b_1 \in \mathbb{Z}_p$ tais que $b_0 \not\equiv b_1 (mod\ p)$.*
Então

$$|(A + b_0) \cup (A + b_1)| \geqslant \alpha + 2.$$

Demonstração. A prova decorre do Teorema 2.2.2 uma vez que se se considerar $A + b_0 = A$ e $k = b_1 - b_0$, obtém-se o enunciado anterior. □

Demonstração. Demonstração do Teorema 2.2.1.
Suponha-se que A é o conjunto com mais elementos, isto é, $|A| \geqslant |B|$. Consequentemente, $\alpha \geqslant \beta$.

Em extensão, os conjuntos A e B representam-se da seguinte forma:

$$A = \{a_0, a_1, \ldots, a_\alpha\} \text{ e } B = \{b_0, b_1, \ldots, b_\beta\}.$$

Considerem-se os seguintes três casos:

(1) Os elementos de A são distintos dos elementos de B, $A \cap B = \emptyset$;

Nos seguintes casos vamos considerar que $A \cap B \neq \emptyset$, vamos dividir esta situação em duas outras: o caso em que o menor conjunto está totalmente contido no maior e o caso onde este está apenas parcialmente contido no maior.

(2) Os elementos de B são iguais a alguns elementos de A, isto é, $B \subset A$;

(3) Alguns elementos de B são iguais a alguns de A mas os restantes são diferentes, isto é, $B \not\subset A$ e $A \cap B \neq \emptyset$.

Em traços gerais esta demonstração consiste em reduzir os casos (1) e (2) ao caso (3) e termina demonstrando-se este último caso. Uma particularidade desta prova é que o caso (1) é reduzido não apenas ao caso (3) mas também ao caso (2).

Caso (1): O problema da determinação da cardinalidade da soma dos conjuntos A e B surge pois, em geral, é difícil contar as repetições de somas que aparecem. Contudo, pretende-se, de acordo com o que foi explicado acima, que de alguma forma a intersecção se torne não vazia. A ideia passa por arranjar outro par de conjuntos em tudo relacionados com os conjuntos originais A e B cuja intersecção seja não vazia.

Seja $k \in \mathbb{Z}$ tal que $A \cap (B - k) \neq \emptyset$.
De forma a compreender-se a trivialidade em encontrar tal k, veja-se, por exemplo, que se $k = b_0 - a_0$ as condições pretendidas são satisfeitas.
No entanto, este processo apenas fará sentido se a cardinalidade de $A + (B - k)$ for igual à cardinalidade de $A + B$. De facto tem-se esta igualdade visto que o conjunto $A + (B - k)$ se refere apenas a uma translação dos elementos de $A + B$. Por conseguinte será suficiente demonstrar o teorema para os conjuntos A e $B - k$. Estes novos conjuntos poderão encaixar-se em qualquer um dos seguintes casos uma vez que, por construção, $|A \cap (B - k)| \neq \emptyset$.

Caso (2): O estudo deste caso será dividido em dois subcasos:

Subcaso (2.1): $A + B = \mathbb{Z}_p$, isto é, $|A + B| = p$.
Neste caso a demonstração fica concluída, uma vez que

$$|A + B| = p \geqslant \alpha + \beta + 1.$$

Subcaso (2.2): $A + B \subsetneq \mathbb{Z}_p$, isto é, $|A + B| < p$.
Considere-se C o subconjunto de $\mathbb{Z}_p$ tal que $C = \mathbb{Z}_p \setminus (A + B)$ e denote-se $|C| = 1 + \gamma \neq 0$.

Neste caso já é sabido que a intersecção dos conjuntos A e B é não vazia mas pretende-se agora que o menor conjunto não esteja contido no maior, para tal, arranjar-se-ão novos conjuntos que dependam dos originais e que verifiquem as condições pretendidas.

Proposição 2.2.1. *Dados dois elementos distintos de B, b_n e b_m, tem-se que*

$$|C \cup (C + b_m - b_n)| \geqslant \gamma + 2.$$

Demonstração. A prova baseia-se no Teorema 2.2.3 quando aplicado ao conjunto C e considerando γ no lugar de α. Pretende-se ver que $\gamma + 2 \leqslant p$. Ora,

$$|A + B| \neq 0, \text{ logo } |A + B| \geqslant 1.$$

$$\text{Como } 1 + \gamma + |A + B| = p \text{ vem que } \gamma + 2 \leqslant p.$$

Para além disso, os elementos de B escolhidos são, por hipótese, distintos por isso

$$b_n \not\equiv b_m (\text{mod } p) \iff 0 \not\equiv b_m - b_n (\text{mod } p).$$

As condições do teorema são assim verificadas, logo

$$|C \cup (C + b_m - b_n)| \geqslant \gamma + 2.$$

Note-se que se pode assumir que B tem dois elementos visto que se $|B| = 1$ o Teorema 2.2.1 é trivial. □

Desta forma e tendo em conta que $|C| = p - |A + B| = \gamma + 1$, podemos afirmar que existe um elemento de $C + b_m - b_n$, x, tal que $x \in A + B$. Ou seja, existem $c_k \in C, a_r \in A$ e $b_s \in B$ tais que

$$c_k + b_m - b_n \equiv a_r + b_s (\text{mod } p).$$

Com o intuito de reduzir este caso ao Caso (3), verifique-se que os conjuntos $A + b_s$ e $B + c_k - b_n$ satisfazem as condições necessárias e, ainda que a demonstração do teorema para estes conjuntos implicará a demonstração para os conjuntos iniciais A e B.

Uma vez que o conjunto $(A + b_s) + (B + c_k - b_n)$ pode ser escrito como $(A + B) + (b_s + c_k - b_n)$ torna-se claro que este novo conjunto constitui apenas uma translação do conjunto $A + B$ e portanto observa-se que

$$|(A + b_s) + (B + c_k - b_n)| = |A + B|.$$

Consequentemente, dar um critério de minimalidade para a cardinalidade da soma destes dois conjuntos será equivalente a mostrá-lo para os conjuntos iniciais A e B.

De forma a serem satisfeitas as hipóteses do Caso (3) resta ver que estes conjuntos têm interseção não vazia e que têm elementos distintos entre si. Tal como visto acima, os conjuntos $A + b_s$ e $B + c_k - b_n$ têm pelo menos um elemento em comum, assim

$$(A + b_s) \cap (B + c_k - b_n) \neq \emptyset.$$

Por outro lado, $B + c_k - b_n$ tem pelo menos um elemento distinto dos elementos de $A + B$, o elemento $b_n + c_k - b_n = c_k \in C$. Encontram-se assim satisfeitas as condições do Caso (3).

Caso (3): Neste caso, a ideia geral será obter, a partir dos conjuntos A e B, conjuntos cuja interseção tenha cardinalidade menor que a cardinalidade da interseção de A com B. A soma dos novos conjuntos é apenas uma translação de $A+B$, condição que garante que basta demonstrar o resultado para os conjuntos transformados. O processo de transformação dos conjuntos repetetir-se-á de modo a que os conjuntos interseção sejam sucessivamente mais pequenos, até se obter o que se designará por conjuntos "limites" e, provando o teorema para estes, concluir-se-á a demonstração. Este é, sem margem para dúvidas, o caso que requer mais engenho tal como se constatará.

Considerem-se os conjuntos $A = \{a_0, \ldots, a_\alpha\}$ e $B = \{b_0, \ldots, b_\beta\}$ tais que $|A \cap B| = 1 + \gamma$ para algum $\gamma \in \mathbb{N}_0$, com $\gamma < \beta$.

Proposição 2.2.2. *Demonstrar o Teorema 2.2.1 para os conjuntos $A \cap B$ e $A \cup B$ será suficiente para demonstrá-lo para os conjuntos A e B.*

Demonstração. Para tal basta ver que todos os elementos de $(A \cap B) + (A \cup B)$ são elementos de $A + B$ e isso é óbvio pelas definições. Desta forma, $(A \cap B) + (A \cup B)$ poderá não ter tantos elementos como $A + B$ mas se $|(A \cap B) + (A \cup B)| \geqslant \alpha + \beta + 1$, então, como $|A + B| \geqslant |(A \cap B) + (A \cup B)|$ sairá o pretendido. □

Assim,

$$|A \cap B| = 1 + \gamma, \text{ e}$$
$$|A \cup B| = \alpha + 1 + \beta + 1 - (1 + \gamma) = 1 + \alpha + \beta - \gamma.$$

Os conjuntos $A \cap B$ e $A \cup B$ estão nas condições do Caso (2). Tal como nesse caso, é possível encontrar conjuntos, digamos $(A \cap B)^*$ e $(A \cup B)^*$, cuja interseção é não vazia e tal que nenhum dos conjuntos está contido no outro.
Designemos $X = (A \cap B)^*$, $Y = (A \cup B)^*$ e $|X \cap Y| = 1 + \delta$, com $\delta \in \mathbb{N}_0 < \gamma$.
Para uma melhor compreensão, note-se em primeiro lugar que $|X| = |A \cap B|$ e que $|Y| = |A \cup B|$ uma vez que a transformação efectuada no caso (2) apenas se refere a uma translação dos conjuntos. No entanto e apesar de X ser obtido através de $A \cap B$, após a transformação que sofreu já não está contido no novo "conjunto união", Y. Deste modo quando se tomar a intersecção de X com Y esta terá menos elementos do que a interseção de A com B e, pelo mesmo motivo, o

conjunto Y será, por seu lado, maior que o conjunto $A \cup B$. Então

$$\begin{aligned} |X \cup Y| &= |X| + |Y| - |X \cap Y| = |A \cap B| + |A \cup B| - |X \cap Y| \\ &= \alpha + \beta + 1 - \gamma + 1 - (1 + \delta) \\ &= 1 + \alpha + \beta - \delta. \end{aligned}$$

Ora, à semelhança do que o que foi feito para os conjuntos A e B, vão-se repetir os mesmos procedimentos mas agora para os conjuntos X e Y.

Utilizando os argumentos do subcaso (2.2), constata-se que provar o Teorema para os conjuntos $X \cap Y$ e $X \cup Y$ será suficiente para prová-lo para os conjuntos A e B. Tem-se ainda que $(X \cap Y) \subseteq (X \cup Y)$ e consequentemente verificam-se de novo nas condições do Caso (2). Repetindo o raciocínio obtém-se uma sucessão de conjuntos "união" e conjuntos "intersecção" que têm respetivamente cardinalidades $1+\alpha+\beta-\gamma$ e $1+\gamma$, $1+\alpha+\beta-\delta$ e $1+\delta$, $1+\alpha+\beta-\epsilon$ e $1+\epsilon$ e assim sucessivamente.
A sucessão de inteiros γ, δ, ϵ,... é decrescente. Como tal, o processo enunciado neste caso terminará quando se obtiver os conjuntos "limite" que terão, respetivamente, cardinalidade $1 + \alpha + \beta$ e 1.

Sejam W e V os conjuntos finais destas sucessivas operações. Pelo que foi argumentado anteriormente, $|W| = 1 + \alpha + \beta$ e $|V| = 1$, denote-se $V = \{v_1\}$. Nestas condições, é trivial que $|W + V| = |W| = 1 + \alpha + \beta$. Fica assim demonstrado o teorema para os conjuntos "limite". O facto de W e V verificarem a afirmação enunciada implica também a prova do teorema para todos os conjuntos anteriores, incluindo para os iniciais A e B. □

Observação 2.2.1. *O argumento utilizado no final da prova da Proposição 2.2.2 baseia-se no facto de que qualquer elemento de $(A \cup B) + (A \cap B)$ é um elemento de $A + B$, logo tem-se que*

$$|A + B| \geqslant |(A \cup B) + (A \cap B)|.$$

Este raciocínio poderia conduzir a interrogações acerca do recíproco. Será que todo o elemento de $A+B$ também é elemento de $(A\cup B)+(A\cap B)$? A resposta será não, tal como se exemplifica de seguida.

Considere-se os seguintes subconjuntos de $\mathbb{Z}_{11}$:

$$A = \{1, 4\} \text{ e } B = \{3, 4\}.$$

Nestas condições tem-se que

$$(A \cup B) + (A \cap B) = \{5, 7, 8\} \text{ e, no entanto, } A + B = \{4, 5, 7, 8\}.$$

2.3. Prova de Davenport do Teorema de Cauchy-Davenport

Em 1935, Davenport apresentou também uma demonstração para o mesmo resultado, tendo por isso ficado conhecido como Teorema de Cauchy-Davenport. A prova original pode ser encontrada em [4].

Teorema 2.3.1. *Sejam $A = \{a_1, \ldots, a_m\}$ um subconjunto de $\mathbb{Z}_p$ com m classes distintas e $B = \{b_1, \ldots, b_n\}$ um subconjunto de $\mathbb{Z}_p$ com n classes distintas, onde p é um número inteiro primo. Sejam $\gamma_1, \ldots, \gamma_l$ todas as classes distintas mod p representáveis como $a_i + b_j$ $(1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n)$.*
Se $m + n - 1 \leqslant p$, então $l \geqslant m + n - 1$ e, caso contrário, $l = p$.

Demonstração. Suponha-se que primeiramente que $m + n - 1 \leqslant p$.

A demonstração será feita por indução no número de elementos de B, n.

Para $n = 1$ nada há a provar, pois nesse caso $l = |A + B| = |A| = m$.

Assim, estude-se o caso em que $n = 2$.
Pretende-se provar que a condição $m + 1 \leqslant p$ implicará $l \geqslant m + 1$. Para tal, suponha-se, com vista a um absurdo, que $m + 1 \leqslant p$ e que $l < m + 1$. Um absurdo será atingido quando se argumentar que $A = \mathbb{Z}_p$, contrariando a hipótese de que $m + 1 \leqslant p$.

Seja A como acima e $B = \{b_1, b_2\}$.

Prove-se que, nestas condições, para cada $a_i \in A$, $a_i + b$ seria um elemento de A.
Ora, $l = |A + B| < m + 1$ (por hipótese) mas, por outro lado, $|A + B| \geqslant |A| = m$. Logo $l = m$.

Como $l = m$ e $|A + b_1| = |A| = m$, todos os elementos de $A + B$ serão obtidos se se calcular, por exemplo, $A + b_1$. Repare-se então que qualquer elemento de $A + b_2$ será uma repetição de algum elemento de $A + b_1$. Desta forma,

$$\forall i \in \{1, \ldots, m\}\ \exists j \neq i \in \{1, \ldots, m\} : a_i + b_1 = a_j + b_2.$$

Defina-se $b = b_2 - b_1 \neq 0$. Logo,

$$\forall i \in \{1, \ldots, m\}\ \exists j \neq i \in \{1, \ldots, m\} : a_i = a_j + b. \tag{2.3}$$

Assim,

$$\forall i \in \{1, \ldots, m\}\ \exists j \neq i \in \{1, \ldots, m\} : a_i + b = a_j + 2b$$

e $a_i + b \in A$ uma vez que a correspondência determinada por 2.3 é bijetiva. Isto é, se a qualquer elemento de A se adicionar a quantidade b obter-se-à outro elemento de A, distinto de a_i. Portanto, para qualquer inteiro u tem-se que $a_i + ub \in A$.

Pelo Lema 2.2.1, $a_i + ub$ representa também as classes de $\mathbb{Z}_p$. Logo viria que $A = \mathbb{Z}_p$, o que é absurdo.

Suponha-se que $n > 2$ e que o mesmo é válido para todo o $n' < n$. Suponha-se também que $l < p$.

Sejam A, B tais como enunciados e designe-se o conjunto dos elementos representáveis como somas de elementos de A e B por C, $C := \{\gamma_1, \ldots, \gamma_l\}$.

Aplicando o teorema aos conjuntos C e $\{b_1, b_n\}$ conclui-se que há pelo menos $l + 1$ classes que são da forma $\gamma_i + b_1$ ou da forma $\gamma_i + b_n$, com $\gamma_i \in C$. Adicionalmente, infere-se que não podem existir apenas classes de uma das formas visto que as classes de cada uma das formas dariam origem a apenas l classes invés das $l + 1$ classes cuja existência é garantida pelo teorema. Assim, existe uma classe δ tal que $\delta - b_1 \in C$ e $\delta - b_n \notin C$.

Considere-se agora todo o conjunto B. Em face do que vimos acima, reordenem-se os elementos de B de modo que o índice r, $1 \leqslant r < n$, separe os elementos de B da seguinte forma:

$$\begin{cases} \gamma_s := \delta - b_s \in C, & 1 \leqslant s \leqslant r \\ \epsilon_t := \delta - b_t \notin C, & r < t \leqslant n. \end{cases}$$

Proposição 2.3.1. *Nenhuma das classes $\gamma_s - b_t$ para $r < t \leqslant n, 1 \leqslant s \leqslant r$ é um elemento do conjunto A.*

Demonstração. Suponhamos que uma destas classes seria um certo $a \in A$. Então, para alguns s e t nas condições acima,

$$\begin{aligned} \gamma_s - b_t = a &\Leftrightarrow a + b_t = \gamma_s \\ &\Leftrightarrow a + b_t = \delta - b_s \\ &\Leftrightarrow a + b_s = \delta - b_t \Leftrightarrow a + b_s = \epsilon_t. \end{aligned}$$

Por um lado, $a + b_s \in C$ mas por outro, $\epsilon_t \notin C$, como $a + b_s = \epsilon_t$ isto é impossível. □

Desta forma, para $1 \leqslant i \leqslant m$ e $r < t \leqslant n$,

$$a_i + b_t \notin \{\gamma_1, \ldots, \gamma_r\}.$$

Ora, visto isto o resultado será simples de concluir.
Construa-se o conjunto $B' = \{b_{r+1}, \ldots, b_n\}$, com cardinalidade

$$n - (r + 1) + 1 = n - r.$$

Uma aplicação do teorema permite retirar que o número de classes representáveis como soma de elementos de A e de B' será maior que uma certa quantidade conveniente, isto é, um número bastante próximo àquele que se prentende alcançar, $m+n-1$. Por outro lado, pelo que foi inferido acima, a cardinalidade da soma dos conjuntos A e B' será naturalmente inferior à cardinalidade do conjunto de elementos representáveis como somas de A e de B, uma vez que não incluem qualquer dos elementos $\gamma_1, \ldots, \gamma_r$.

Tomando $l' = |A + B'|$, resulta então o seguinte enquadramento

$$m + n - r - 1 \leqslant l' \leqslant l - r,$$

de onde se conclui que

$$l \geqslant m + n - 1.$$

De forma a concluir a demonstração resta provar o último caso do teorema em questão, que será feita com recurso a uma redução dos conjuntos inicias. Repare-se que para estudar o caso em que $m + n - 1 > p$ será suficiente considerar os conjuntos $A' = A$ e $B' := \{b_1, \ldots, b_{p+1-m}\}$.

De acordo com o acima estipulado, tem-se que $A + B \supseteq A' + B'$. Assim se se provar que $A' + B' = \mathbb{Z}_p$ a demonstração ficará concluída.

Ora

$$|A'| + |B'| - 1 = m + p + 1 - m - 1 = p,$$

e portanto é possível aplicar o teorema, donde resulta que $|A' + B'| \geqslant |A'| + |B'| - 1 = p$. Logo $A' + B' = \mathbb{Z}_p$ e, consequentemente, $A + B = \mathbb{Z}_p$. □

3. Prova da conjetura de Erdős e Heilbronn

No presente capítulo serão apresentados os resultados centrais do artigo [7], o qual se pode dizer que é o culminar de toda a teoria anteriormente apresentada, uma vez que contém a primeira demonstração para a Conjetura de Erdős e Heilbronn. Em virtude do desenvolvimento e aprofundamento das técnicas de Álgebra Multilinear na resolução de problemas da teoria aditiva, foi possível quebrar a conjetura em questão. Sem fugir à regra, estes avanços possibilitaram o aparecimento de muitas outras técnicas que demonstram de forma mais simples, os resultados apresentados nesta dissertação. Como exemplo, expõe-se o Método Polinomial, apresentado em [1], desenvolvido por Noga Alon, Melvyn B. Nathanson e Imre Ruzsa. Mais desenvolvimentos deste método podem ser consultados em [16].

3.1. Subespaço Cíclico para as Derivadas de Grassmann

No artigo [7], dá-se especial atenção ao operador Df uma vez que o seu espetro se relaciona com as somas de m elementos de um conjunto que, tal como foi possível comprovar até então, estão no centro desta dissertação.

Ao longo desta secção, considere-se $\mathbb{F}$ um corpo de caraterística p e V um espaço vetorial sobre $\mathbb{F}$ de dimensão finita d.

Nesta secção a definição de $Q_{m,d}$ será ligeiramente diferente, tal como adotado em [7]. Neste caso as aplicações de $Q_{m,d}$ serão as aplicações crescentes de $\{0,\ldots,m-1\}$ para $\{0,\ldots,d-1\}$.

Recorde-se que dado um operador linear f em V, a sua *derivada* de Grassmann em $\bigwedge^m V$ define-se como

$$Df(v_1 \wedge \ldots \wedge v_m) = \sum_{i=1}^{m} v_1 \wedge \ldots \wedge v_{i-1} \wedge f(v_i) \wedge v_{i+1} \wedge \ldots \wedge v_m,$$

e ainda que, dado um subconjunto A de $\mathbb{F}$, $\bigwedge^m A$ denota o conjunto das somas de todos os subconjuntos de A com cardinalidade m.

Proposição 3.1.1. *Seja $f \in L(V,V)$ e $\sigma(f) = \{\lambda_0, \ldots, \lambda_{d-1}\}$. Tem-se que*

$$\sigma(Df) = \left\{ \sum_{j=0}^{m-1} \lambda_{\alpha(j)} : \alpha \in Q_{m,d} \right\} = \bigwedge{}^m \sigma(f).$$

Demonstração. A prova decorre de se aplicar a técnica da demonstração da Proposição 1.4.8, isto é, considerar uma base de vetores próprios de f. □

Seja V um espaço vetorial tal como acima referido e seja $\{e_0, \ldots, e_{d-1}\}$ uma base de V. Recordando a Proposição 1.4.6, o conjunto

$$\varrho = \{e_{\alpha(0)} \wedge \ldots \wedge e_{\alpha(m-1)} : \alpha \in Q_{m,d}\},$$

é uma base de $\wedge^m V$.
Seja λ uma partição de comprimento m, tal que $\lambda_m \leqslant d - m$. Ora, pela definição de partição, $\lambda_1 \leqslant \ldots \leqslant \lambda_m$ e, consequentemente,

$$\lambda_1, \lambda_2 + 1, \ldots \ , \lambda_m + m - 1 \tag{3.1}$$

é uma sucessão estritamente crescente. Assim, $\lambda_1, \lambda_2 + 1, \ldots \ , \lambda_m + m - 1$ é uma aplicação de $Q_{m,d}$, uma vez que a condição acima imposta, $\lambda_m \leqslant d - m$, garante que $\lambda_m + m - 1 \leqslant d - 1$. Reciprocamente, tomando uma aplicação α de $Q_{m,d}$ obtém-se a seguinte sucessão crescente no sentido lato,

$$\alpha_1 \leqslant \alpha_2 - 1 \leqslant \ldots \leqslant \alpha_m - (m-1).$$

Sendo isto uma bijeção, todas as aplicações de $Q_{m,n}$ se escrevem na forma 3.1. Consequentemente, tem-se

$$\varrho = \{e_{\alpha(0+\lambda_1)} \wedge \ldots \wedge e_{\alpha(m-1+\lambda_m)} : \lambda \in \mathscr{P}_m \text{ e } \lambda_m \leqslant d - m\}.$$

De seguida apresentam-se resultados cujas demonstrações que podem ser encontradas no artigo [7].

Teorema 3.1.1. *Seja λ uma partição de k. Então*

$$\xi_\lambda(id) = \frac{k!}{\prod_{ij} h^\lambda_{i,j}}.$$

Nos teoremas seguintes, S_k será identificado com o estabilizador de $k+1$ em S_{k+1}. Relembre-se que o estabilizador de $k+1$ consiste no subgrupo de permutações que mantêm este elemento fixo.

Teorema 3.1.2. *Seja λ uma partição de $\mathscr{P}_{k+1,m}$. Então*

$$\xi_\lambda = \sum_{\mu \in \mathscr{P}_{k,m},\ \mu \to \lambda} \xi_\mu.$$

Pelos Teoremas acima apresentados, sem demonstração, infere-se diretamente o seguinte corolário.

Corolário 3.1.1. *Seja $\lambda \in \mathscr{P}_{k+1,m}$. Então*

$$\frac{k+1}{\prod_{ij} h^\lambda_{i,j}} = \sum_{\mu \to \lambda} \frac{1}{\prod_{ij} h^\lambda_{i,j}}.$$

Demonstração. Para $k = 0$ o resultado é obviamente verdadeiro. Suponha-se $k > 0$. Então pelos Teoremas 3.1.1 e 3.1.2,

$$\frac{(k+1)!}{\prod_{ij} h^\lambda_{i,j}} = \xi_\lambda(\mathrm{id}) = \sum_{\mu \to \lambda} \frac{k!}{\prod_{ij} h^\lambda_{i,j}}.$$

□

Definição 3.1.1. *Sejam $f \in L(V,V)$ e $\lambda \in \mathscr{P}_m$, define-se:*

$$\bigwedge(f^\lambda)(v) = f^{\lambda_1}(v) \wedge f^{\lambda_2+1}(v) \wedge \ldots \wedge f^{\lambda_m+m-1}(v).$$

Teorema 3.1.3. *Seja V um espaço vetorial sobre $\mathbb{F}$ e seja $f \in L(V,V)$. Então:*

$$(Df)^k(v \wedge f(v) \wedge \ldots. \wedge f^{m-1}(v)) = \sum_{\lambda \in \mathscr{P}_{k,m}} \frac{k!}{\prod_{ij} h^\lambda_{i,j}} \bigwedge(f^\lambda)(v).$$

Demonstração. A prova será feita por indução em k.

Se $k = 0$,

$$(Df)^0(v \wedge f(v) \wedge \ldots. \wedge f^{m-1}(v)) = (v \wedge f(v) \wedge \ldots. \wedge f^{m-1}(v))$$

e, convencionando, que $\prod_\emptyset h^\lambda_{i,j} = 1$ para a partição nula, virá que

$$\sum_{\lambda \in \mathscr{P}_{0,m}} \frac{0!}{\prod_\emptyset h^\lambda_{i,j}} \bigwedge(f^\lambda)(v) = \sum_{\lambda \in \mathscr{P}_{0,m}} \frac{0!}{1} \bigwedge(f^\lambda)(v) = \bigwedge(f^0)(v) = (v \wedge f(v) \wedge \ldots. \wedge f^{m-1}(v)).$$

A última igualdade acontece uma vez que a única partição nestas condições é a partição nula e portanto não afecta em nada os expoentes do tensor antissimétrico.

Suponha-se que o resultado é válido para k e prove-se de seguida para $k+1$. Deste modo

tem-se

$$(Df)^k(v \wedge f(v) \wedge \ldots . \wedge f^{m-1}(v)) = \sum_{\lambda \in \mathscr{P}_{k,m}} \frac{k!}{\prod_{ij} h^{\lambda}_{i,j}} \bigwedge (f^{\lambda})(v).$$

Logo,

$$(Df)^{k+1}(v \wedge f(v) \wedge \ldots . \wedge f^{m-1}(v)) = (Df)\bigg((Df)^k(v \wedge f(v) \wedge \ldots . \wedge f^{m-1}(v))\bigg)$$
$$= (Df)\bigg(\sum_{\lambda \in \mathscr{P}_{k,m}} \frac{k!}{\prod_{ij} h^{\lambda}_{i,j}} \bigwedge (f^{\lambda})(v)\bigg).$$

Utilizando a linearidade do operador Df e a Definição 3.1.1, vem que

$$(Df)^{k+1}(v \wedge f(v) \wedge \ldots . \wedge f^{m-1}(v)) = \sum_{\lambda \in \mathscr{P}_{k,m}} \frac{k!}{\prod_{ij} h^{\lambda}_{i,j}} (Df) \bigwedge (f^{\lambda})(v)$$
$$= \sum_{\lambda \in \mathscr{P}_{k,m}} \frac{k!}{\prod_{ij} h^{\lambda}_{i,j}} (Df)(f^{\lambda_1}(v) \wedge f^{\lambda_2+1}(v) \wedge \ldots \wedge f^{\lambda_m+m-1}(v)).$$

Aplicando a definição de derivada, Definição 1.4.3, vem que

$$(Df)^{k+1}(v \wedge f(v) \wedge \ldots . \wedge f^{m-1}(v)) =$$
$$= \sum_{\lambda \in \mathscr{P}_{k,m}} \frac{k!}{\prod_{ij} h^{\lambda}_{i,j}} \sum_{t=1}^{m} (f^{\lambda_1}(v) \wedge f^{\lambda_2+1}(v) \wedge \ldots \wedge f\big(f^{\lambda_t+t-1}(v)\big) \wedge \ldots \wedge f^{\lambda_m+m-1}(v))$$
$$= \sum_{\lambda \in \mathscr{P}_{k,m}} \frac{k!}{\prod_{ij} h^{\lambda}_{i,j}} \sum_{t=1}^{m} (f^{\lambda_1}(v) \wedge f^{\lambda_2+1}(v) \wedge \ldots \wedge f^{\lambda_t+t}(v) \wedge \ldots \wedge f^{\lambda_m+m-1}(v))$$
$$= \sum_{\lambda \in \mathscr{P}_{k,m}} \frac{k!}{\prod_{ij} h^{\lambda}_{i,j}} \sum_{t=1}^{m} \bigwedge_{i=1}^{m} f^{\lambda_i+(i-1)+\delta_{i,t}}(v) = \sum_{\lambda \in \mathscr{P}_{k,m}} \frac{k!}{\prod_{ij} h^{\lambda}_{i,j}} \sum_{t=0}^{m-1} \bigwedge_{i=0}^{m-1} f^{\lambda_{i+1}+i+\delta_{i,t}}(v).$$

Consideremos de seguida uma propriedade das aplicações multilineares que aliás já foi enunciada.

Propriedade 3.1.1. *Dados índices inteiros positivos i e j, com $i \neq j$, tais que $x_i = x_j$ então, $x_1 \wedge x_2 \wedge \ldots \wedge x_m = 0$.*

Assim, podemos retirar as parcelas que incluem repetições nas componentes das partições

logo, tem-se que:

$$(Df)^{k+1}(v \wedge f(v) \wedge \ldots . \wedge f^{m-1}(v)) = \sum_{\lambda \in \mathscr{P}_{k,m}} \frac{k!}{\prod_{ij} h_{i,j}^{\lambda}} \sum_{t \notin \{i:\lambda_i = \lambda_{i+1}\}} \bigwedge_{i=0}^{m-1} f^{\lambda_{i+1}+i+\delta_{i,t}}(v)$$

$$= \sum_{\lambda \in \mathscr{P}_{k,m}} \frac{k!}{\prod_{ij} h_{i,j}^{\lambda}} \sum_{\lambda \to \chi} \bigwedge (f^{\chi})(v) = \sum_{\chi \in \mathscr{P}_{k+1,m}} \left(\sum_{\lambda \to \chi} \frac{k!}{\prod_{ij} h_{i,j}^{\chi}} \right) \bigwedge (f^{\chi})(v)$$

$$= \sum_{\chi \in \mathscr{P}_{k+1,m}} \frac{(k+1)k!}{\prod_{ij} h_{i,j}^{\chi}} \bigwedge (f^{\chi})(v) = \sum_{\chi \in \mathscr{P}_{k+1,m}} \frac{(k+1)!}{\prod_{ij} h_{i,j}^{\chi}} \bigwedge (f^{\chi})(v).$$

□

Observação 3.1.1. *Note-se que a troca de somatórios é válida pois esta apenas implica uma reordenação das parcelas. Para além disso, pela definição de uma partição seguir de outra, sabe-se que as partições em questão apenas diferem uma unidade numa componente, visto que $\lambda \in P_{k,m}$ e que $\lambda \to \chi$ segue então que $\chi \in P_{k+1,m}$. A penúltima igualdade deve-se ao Corolário 3.1.1.*

Apresenta-se de seguida, o último teorema necessário para a demonstração da conjetura estabelecida por Erdős e Heilbronn.

Seja g um operador linear sobre um espaço vetorial V. Para cada $x \in V$, denota-se $\varphi_g(x)$ o subespaço cíclico $\langle g^i(x) : i \geqslant 0 \rangle$. É também necessário definir uma ordem no conjunto $\mathscr{P}_m$, sejam os elementos ordenados primeiramente por grau e depois, dentro dos do mesmo grau, por ordem lexicográfica. Aplique-se esta ordem total, que se denotará por $\leqslant$, ao conjunto ϱ.

Teorema 3.1.4. *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$ e $v \in V$. Se f é um operador linear em V, então*

$$\dim \varphi_{Df}(v \wedge f(v) \wedge \ldots . \wedge f^{m-1}(v)) \geqslant \min\{p, (\dim \varphi_f(v) - m)m + 1\}.$$

Demonstração. Seja $n = \dim \varphi_f(v) = \dim < f^i(v) : i \geqslant 0 >$.

Pretende-se ver que
$\dim \varphi_{Df}(v \wedge f(v) \wedge \ldots \wedge f^{m-1}(v)) \geqslant \min\{p, (n-m)m + 1\}$. Observe-se que faz sentido escrever $n - m$, uma vez que se $m > n$ haveria outro representante para m que fosse menor do que n.

Seja $k \leqslant \min\{p - 1, m(n - m)\}$.

A ideia da demonstração será mostrar que a família $(Df)^k(v \wedge f(v) \wedge \ldots \wedge f^{m-1}(v))$ ainda é uma família linearmente independente e, consequentemente, o grau do subespaço cíclico terá de ser naturalmente maior que k, isto é, maior ou igual a $\min\{p, (n-m)m + 1\}$.

Faça-se a divisão inteira de k por m. Assim, existem r e t inteiros não negativos tais que $0 \leqslant t < m$ e $k = mt + r$. Seja ω a seguinte partição de grau k:

$$\omega = \big(\underbrace{t, \ldots , t}_{\text{m-r vezes}}, \underbrace{t+1, \ldots , t+1}_{\text{r vezes}}\big).$$

Pelo Teorema anterior sabe-se que

$$(Df)^k(v \wedge f(v) \wedge \ldots . \wedge f^{m-1}(v)) = \sum_{\lambda \in \mathscr{P}_{k,m}} \frac{k!}{\prod_{ij} h^{\lambda}_{i,j}} \bigwedge (f^{\lambda})(v). \tag{3.2}$$

Pretende-se mostrar que é possível escrever $(Df)^k(v \wedge f(v) \wedge \ldots . \wedge f^{m-1}(v))$ como uma soma de parcelas com grau exactamente igual ao grau da partição ω e parcelas com grau estritamente menor que o grau de ω. Deste modo, poderá argumentar-se que a família $(Df)^k(v \wedge f(v) \wedge \ldots . \wedge f^{m-1}(v))$ ainda é uma família linearmente independente fazendo alusão à Proposição 2.1.1.

Para tal, comece-se por verificar que $\bigwedge(f^{\omega})(v)$ é o maior elemento de (3.2) no qual $(Df)^k(v \wedge f(v) \wedge \ldots . \wedge f^{m-1}(v))$ tem um coeficiente não nulo.
Provemos que

$$\frac{k!}{\prod_{ij} h^{\omega}_{i,j}} \text{ é um coeficiente não nulo.}$$

Ora, esta fração é nula se $k! \equiv 0 \pmod{p}$. Porém, $k < p$ e portanto tem-se que $p \nmid k!$. Assim, $k! \not\equiv 0 \pmod{\text{p}}$ e deste modo os coeficientes de $\bigwedge(f^{\omega})(v)$ são inteiros não nulos módulo p.

Veja-se agora que $\bigwedge(f^{\omega})(v)$ é o maior elemento de (3.2).

A ideia será encontrar uma expressão que seja igual a (3.2) na qual os expoentes dos operadores aparecem escritos da forma mais simples possível, isto é, escrevendo-os à custa de expoentes menores que n. Esta é a simplificação pretendida uma vez que, para expoentes maiores ou iguais a n, como $n = \dim \varphi_f(x)$, estes poderão ser escritos à custa dos primeiros. Tal como foi feito na demonstração do Teorema 2.1.1, iremos recorrer à denominação informal de expoente "grande" e "pequeno". Neste caso, dada uma partição λ, diremos que o expoente do operador f é "grande" quando a maior potência (isto é, a última) do desenvolvimento de $\bigwedge(f^{\omega})$ for igual ou superior à dimensão do subespaço cíclico $\varphi_f(x)$, n, caso contrário, diremos que f tem expoente pequeno.

Assim, qual deverá ser a divisão em casos que corresponderá a uma análise de acordo com o "tamanho" do expoente de f? De acordo com o que foi dito anteriormente e, recordando a Definição 3.1.1, o expoente de f será pequeno se $\lambda_m + m - 1 < n$.

Compreendendo agora em pleno a escolha dos diferentes casos, estamos então em condições de proceder ao seu estudo.
Seja $\lambda = (\lambda_1, \ldots, \lambda_m)$ uma partição arbitrária de $\mathscr{P}_{k,m}$.

Caso (i): Suponha-se que f tem expoente "pequeno", isto é, $\lambda_m < n - m + 1$.

As partições λ e ω pertencem a $\mathscr{P}_{k,m}$ logo têm o mesmo grau e para compreender qual das partições será maior será necessário considerar a ordem lexicográfica. De facto, ω é a maior partição em $\mathscr{P}_{k,m}$. Se existisse alguma partição, em $\mathscr{P}_{k,m}$, maior que ω esta teria de satisfazer alguma das seguintes condições:

- numa das primeiras $(m - r)$ componentes teria de aparecer o inteiro $t + 1$

- ser composta por inteiros maiores que $t+1$ a partir de alguma das r últimas componentes.

No entanto, qualquer um desses casos implicaria que a partição em questão não teria grau k logo, a maior partição de $\mathscr{P}_{k,m}$ é ω. Note-se que este argumento baseia-se fundamentalmente no algortimo da divisão. Consequentemente, como a ordem de $\mathscr{P}_m$ se estende à ordem de ϱ, conclui-se que

$$\forall \lambda \in \mathscr{P}_{k,m}, \quad \bigwedge(f^{\lambda})(v) < \bigwedge(f^{\omega})(v),$$

permitindo assim reduzir uma parte de (3.2), a parcelas com grau estritamente menor que o grau de ω.

Caso (ii): Admita-se agora se que $\lambda_m \geqslant n-m+1$, isto é, que f tem expoente "grande".

Assim, visto que $\dim \varphi_f(v) = n$, existem $x_1, \ldots, x_n \in \mathbb{F}$ tais que $f^{\lambda_m+m-1}(v) = x_1 v + \ldots + x_n f^{n-1}(v)$. Nestas condições tem-se então que $\bigwedge(f^{\lambda})(v)$ é combinação linear de elementos de ϱ com grau inferior ao grau de λ. Como

$$\bigwedge(f^{\lambda})(v) = f^{\lambda_1}(v) \wedge f^{\lambda_2+1}(v) \wedge \ldots \wedge x_1 v + \ldots + x_n f^{n-1}(v) =$$

$$= \underbrace{f^{\lambda_1}(v) \wedge f^{\lambda_2+1}(v) \wedge \ldots \wedge x_1 v}_{(1)} + \ldots + \underbrace{f^{\lambda_1}(v) \wedge f^{\lambda_2+1}(v) \wedge \ldots \wedge x_n f^{n-1}(v)}_{(n)},$$

bastará ver que cada parcela (i) tem grau inferior ao de λ, com $i = 1, 2, \ldots, n$. Para tal, a soma dos expoentes de (i) deverá ser menor que a soma dos expoentes de $\bigwedge(f^{\lambda})(v)$, pois

$$\deg(\lambda) = (\lambda_1 + \lambda_2 + 1 + \ldots + \lambda_m + m - 1) - (1 + 2 + \ldots + (m-1))$$

$$\text{e } \deg(i) = (\lambda_1 + \lambda_2 + 1 + \ldots + (i-1)).$$

Mas, na realidade, é suficiente verificá-lo para $i = n$, uma vez que esta é a parcela com soma máxima dos seus expoentes. Quer-se então ver que

$$\lambda_m > n - 1.$$

De acordo com a suposição do Caso (ii), $\lambda_m \geqslant n - m + 1$. No pior caso $m = 1$ e aí $\lambda_m \geqslant n$ e claramente obtém-se que $\lambda_m > n - 1$.

Após analisar os casos anteriores, é então possível escrever
$(Df)^k(v \wedge f(v) \wedge \ldots. \wedge f^{m-1}(v))$ como

$$(Df)^k(v \wedge f(v) \wedge \ldots. \wedge f^{m-1}(v)) = \frac{k!}{\prod_{ij} h_{i,j}^{\omega}} \bigwedge(f^{\omega})(v) + \sum_{\lambda<\omega} \bigwedge(f^{\lambda})(v),$$

onde a primeira parcela diz respeito ao maior elemento de ϱ e a segunda se refere os elementos

que têm grau inferior ao grau de ω. Esta segunda parcela contém, para além dos elementos de ϱ que já tinham grau inferior ao de ω, os elementos cujos graus foram reescritos, de acordo com o que foi visto no Caso (ii), de forma a ficarem também inferiores ao de ω.

Pela Proposição 2.1.1, tem-se agora que
$\{(Df)^k(v \wedge f(v) \wedge \ldots \wedge f^{m-1}(v)) : 0 \leqslant k \leqslant \min\{p-1, (n-m)m\}\}$ é um conjunto linearmente independente e, consequentemente,

$$\dim \varphi_{Df}(v \wedge f(v) \wedge \ldots \wedge f^{m-1}(v)) \geqslant \min\{p, (\dim \varphi_f(v) - m)m + 1\}.$$

□

Corolário 3.1.2. *Seja f um operador linear em V. Então,*

$$\deg(P_{Df}) \geqslant \min\{p, (\deg(P_f) - m)m + 1\}.$$

Demonstração. Por [13], Teorema 6, p.397, tem-se que a dimensão máxima dos subespaços cíclicos de f em V é igual ao grau do polinómio mínimo de f. Logo, o Corolário segue imediatamente do Teorema 3.1.4. □

Teorema 3.1.5. *Seja A um subconjunto finito de um corpo $\mathbb{F}$ e m um inteiro positivo. Então*

$$|\bigwedge^m A| \geqslant \min\{p, m|A| - m^2 + 1\}.$$

Demonstração. Seja f um operador linear diagonal em $\mathbb{F}^n$ tal que $\sigma(f) = A$. Sabe-se, da Proposição 3.1.1 que

$$\sigma(Df) = \bigwedge^m A.$$

Uma vez que Df é diagonal, $\deg(P_{Df})$ é igual ao número de valores próprios distintos de Df. Portanto, usando o Corolário 3.1.2,

$$|\bigwedge^m A| = |\sigma(Df)| = \deg(P_{Df}) \geqslant \min\{p, m|A| - m^2 + 1\}.$$

□

Em particular, para $A \subseteq \mathbb{Z}_p$, $|A \wedge A| \geqslant \min\{p, 2|A| - 3\}$. Deste modo obtém-se então uma demonstração para a conjetura de Erdős e Heilbronn, resolvendo assim uma conjetura com aproximadamente 30 anos.

O exemplo seguinte permite observar que a estimativa encontrada no Teorema 3.1.5 se verifica com igualdade, justificando que este é o menor limite inferior possível.
Sejam $F = \mathbb{Z}_p$ e A um conjunto cujos elementos formam uma progressão aritmética.

Exemplo 3.1.1. *Sejam $a \in \mathbb{N}$, com $a \leqslant p$ e $S = \{1, \ldots, a\}$. Considere-se agora o epimorfismo canónico $\nu : \mathbb{Z} \to \mathbb{Z}_p$ e seja $A = \nu(S)$. Uma vez que a é inferior a p, tem-se que as classes ($\mod p$) são também todas distintas e $|A| = |S| = a$.*

Pretende-se calcular $|\bigwedge^m A|$ e ver que se justifica a possível igualdade do Teorema 3.1.5. Ora, $|\bigwedge^m A| = |\bigwedge^m S|$ e o primeiro elemento de $\bigwedge^m S$ será obtido através da soma dos primeiros m elementos, isto é, $1 + \ldots + m = \frac{m(m+1)}{2}$. Naturalmente, o último elemento será a soma dos m últimos elementos de S. Logo, será dado por

$$a - (m-1) + \ldots + a = \frac{(2a - (m-1))m}{2} = ma - \frac{m(m-1)}{2}.$$

Assim,

$$\bigwedge^m S = \left[\frac{m(m+1)}{2}, ma - \frac{m(m-1)}{2}\right].$$

Tem-se então que

$$\begin{aligned} |\bigwedge^m A| = |\bigwedge^m S| &= ma - \frac{m(m-1)}{2} - \frac{m(m+1)}{2} + 1 \\ &= ma - \frac{m^2 - m + m^2 + m}{2} + 1 = ma - m^2 + 1 = m(|A| - m) + 1, \end{aligned}$$

exemplificando deste modo que a estimativa anterior é, de facto, a melhor possível para $\mathbb{Z}_p$. Note-se que ainda teria de ser calculado o mínimo entre p e $m(|A| - m) + 1$, no entanto, escolhendo um p suficientemente grande de forma a que o mínimo seja $m(|A| - m) + 1$, o exemplo continua a ser válido e serve o propósito.

3.2. Problemas Aditivos

Nesta última secção pretende-se exemplificar um conjunto de resultados extraordinários aos já apresentados que podem ser demonstrados usando as técnicas explicitadas anteriormente e que podem ser encontrados em [16] e [7]. Estes referem-se à obtenção de $\mathbb{Z}_p$ como conjunto de somas de certos conjuntos.

Teorema 3.2.1. *Seja $A \subseteq \mathbb{Z}_p$ com cardinalidade $\lfloor\sqrt{4p-7}\rfloor + 1$. Então, todo o elemento de $\mathbb{Z}_p$ é a soma de um subconjunto de A de cardinalidade $\lfloor(\lfloor\sqrt{4p-7}\rfloor + 1)/2\rfloor$.*

Demonstração. Seja $m = \left\lfloor \frac{|A|}{2} \right\rfloor$. Pelo Teorema 3.1.5,

$$|\bigwedge{}^m A| \geqslant \min\left\{p, \left\lfloor \frac{|A|}{2} \right\rfloor |A| - \left(\left\lfloor \frac{|A|}{2} \right\rfloor\right)^2 + 1\right\}$$

$$= \min\left\{p, \left\lfloor \frac{|A|^2}{4} \right\rfloor + 1\right\} \geqslant \min\{p, (p-1)+1\} = p.$$

Logo $|\bigwedge^m A| = \mathbb{Z}_p$.

Resta então provar que $\left\lfloor \frac{|A|}{2} \right\rfloor |A| - \left(\left\lfloor \frac{|A|}{2} \right\rfloor\right)^2 = \left\lfloor \frac{|A|^2}{4} \right\rfloor$ e que $\left\lfloor \frac{|A|^2}{4} \right\rfloor \geqslant (p-1)$.

Para verificar a igualdade basta apenas considerar o caso em que $|A|$ é um inteiro ímpar pois, caso contrário $\left\lfloor \frac{|A|}{2} \right\rfloor$ é apenas $\frac{|A|}{2}$ e este não levanta qualquer dúvida. Portanto, quando $|A| = 2k+1$, para algum $k \in \mathbb{Z}$,

$$\left\lfloor \frac{|A|}{2} \right\rfloor = \left\lfloor \frac{2k+1}{2} \right\rfloor = \left\lfloor k + \frac{1}{2} \right\rfloor = k,$$

logo,

$$\left\lfloor \frac{|A|}{2} \right\rfloor |A| - \left(\left\lfloor \frac{|A|}{2} \right\rfloor\right)^2 + 1 = k\dot{(}2k+1) - k^2 = k^2 + k.$$

Por outro lado,

$$\left\lfloor \frac{|A|^2}{4} \right\rfloor = \left\lfloor \frac{(2k+1)^2}{4} \right\rfloor = \left\lfloor k^2 + k + \frac{1}{4} \right\rfloor = k^2 + k.$$

Para verificar a desigualdade note-se que $\lfloor\sqrt{4p-7}\rfloor + 1 = \lceil\sqrt{4p-4}\ \rceil$. Ora, uma vez que entre $4p-7$ e $4p-4$ têm de estar todos os representantes de $\mathbb{Z}_4$ e que os quadrados perfeitos têm de ser congruentes com 0 ou 1 em $(mod\ 4)$, compreende-se que $4p-7$ e $4p-4$ são os únicos que poderão originar raízes inteiras e, ainda mais, consecutivas. Assim, $\lfloor\sqrt{4p-7}\rfloor + 1 = \lceil\sqrt{4p-4}\ \rceil$. Logo,

$$\left\lfloor \frac{|A|^2}{4} \right\rfloor = \left\lfloor \frac{(\lfloor\sqrt{4p-7}\rfloor + 1)^2}{4} \right\rfloor = \left\lfloor \frac{(\lceil\sqrt{4p-4}\ \rceil)^2}{4} \right\rfloor \geqslant \frac{4p-4}{4} = p-1.$$

□

Observação 3.2.1. *Note-se que a cardinalidade tomada no Teorema 3.2.1 é a melhor possível pois, se se tomar um conjunto A tal como no Exemplo 3.1.1, tem-se que $|\bigwedge^m A| = m(|A|-m)+1$*

e, se se permitir $|A| \leqslant \lfloor\sqrt{4p-7}\rfloor$*, obtém-se* $|\bigwedge^m A| \leqslant p-1-3/4 \leqslant p-1$*, evidenciando assim a existência de elementos de* $\mathbb{Z}_p$ *que não são somas de subconjuntos de* A *de cardinalidade* m*.*

No resultado anterior, não foram permitidos conjuntos com cardinalidade diferente de $\left\lfloor\frac{|A|}{2}\right\rfloor$ mas foi permitido o caso em que $0 \in A$. No próximo resultado, toma-se uma abordagem ligeiramente diferente, excluindo o caso em que $0 \in A$ e permitindo subconjuntos de cardinalidade $\left\lfloor\frac{|A|-1}{2}\right\rfloor$ e $\left\lfloor\frac{|A|+1}{2}\right\rfloor$.

Corolário 3.2.1. *Seja* $A \subseteq \mathbb{Z}_p \setminus \{0\}$ *tal que* $|A| = \lfloor\sqrt{4p-7}\rfloor$*. Então, todo o elemento de* $\mathbb{Z}_p$ *é soma de um subconjunto de* A *com cardinalidade* $\left\lfloor\frac{|A|-1}{2}\right\rfloor$ *ou* $\left\lfloor\frac{|A|+1}{2}\right\rfloor$*.*

Demonstração. Seja $S = A \cup \{0\}$. Pelo Teorema 3.2.1, todo o elemento de $\mathbb{Z}_p$ é soma de um subconjunto de A de cardinalidade $\left\lfloor\frac{|S|}{2}\right\rfloor$. Um subconjunto de S que contém o 0 corresponde a um subconjunto de A com cardinalidade $\left\lfloor\frac{|S|}{2}\right\rfloor - 1 = \left\lfloor\frac{|A|+1}{2}\right\rfloor - 1 = \left\lfloor\frac{|A|+1-2}{2}\right\rfloor = \left\lfloor\frac{|A|-1}{2}\right\rfloor$ e, um subconjunto de S que não contém o 0 é um subconjunto de A de cardinalidade $\left\lfloor\frac{|S|}{2}\right\rfloor = \left\lfloor\frac{|A|+1}{2}\right\rfloor$. □

Observação 3.2.2. *O Teorema 3.2.1 mantém-se válido se se substituir* $\lfloor(\lfloor\sqrt{4p-7}\rfloor+1)/2\rfloor$ *por* $\lceil(\lfloor\sqrt{4p-7}\rfloor+1)/2\rceil$*, verifique-se:*
Com $m = \lceil(\lfloor\sqrt{4p-7}\rfloor+1)/2\rceil$*, tem-se ainda que* $m(|A|-m) = \left\lfloor\frac{|A|^2}{4}\right\rfloor$ *e, portanto o teorema continua válido. Se se fizer uma substituição análoga no Corolário 3.2.1 o resultado também se mantém.*

Verifique-se que se $\lfloor\sqrt{4p-7}\rfloor$ for um número ímpar então, a limitação $\lfloor\sqrt{4p-7}\rfloor$ no Corolário 3.2.1 é a melhor possível. Este exemplo foi estabelecido por Erdős e Heilbronn.

Exemplo 3.2.1. *Seja* $A = \{-s, \ldots, -1\} \cup \{1, \ldots, s\}$*, onde* $2s+1 = \lfloor\sqrt{4p-7}\rfloor$*.*

A conclusão será imediata se se provar que o conjunto das somas de todos os subconjuntos de A*, mod* p*, tem cardinalidade* $1+s^2+s \leqslant (p-1)$*. Desta forma, vemos que* $|A|$ *tem cardinalidade* $\lfloor\sqrt{4p-7}\rfloor - 1$ *mas, no entanto, as somas de subconjuntos de* A *de qualquer cardinalidade não cobrem todo o espaço* $\mathbb{Z}_p$ *portanto, as somas de cardinalidade* $\left\lfloor\frac{|A|-1}{2}\right\rfloor$ *ou* $\left\lfloor\frac{|A|+1}{2}\right\rfloor$ *também não irão, de certeza, cobrir* $\mathbb{Z}_p$*.*

Ora, o conjunto das somas de todos os subconjuntos tem a cardinalidade acima indicada uma vez que o conjunto vai conter todos os números, positivos e negativos, entre 1 *e* $\frac{s(s+1)}{2}$*, portanto tem precisamente* $\frac{s(s+1)}{2} \times 2 + 1$*, onde a adição de um elemento representa todas as somas possíveis que conduzem a* 0*, o que termina o exemplo.*

3.3. Método Polinomial

Em 1995, Noga Alon, Melvyn B. Nathanson e Imre Ruzsa produziram um artigo no qual podemos encontrar um método simplificado para a demonstração da Conjetura de Erdős e Heilbronn envolvendo apenas argumentos polinomiais. De facto, alguns dos argumentos utilizados lembram o tratamento feito com Álgebra Multilinear, tal como se poderá confirmar.

O método polinomial apresentado no artigo [1] tem como base de partida dois lemas que se apresentam de seguida, bem como as suas demonstrações.

Lema 3.3.1. *Sejam A e B subconjuntos não vazios de um corpo $\mathbb{F}$ tais que $|A| = k$ e $|B| = l$. Seja $f(x, y)$ um polinómio com coeficientes no corpo e com grau no máximo $k-1$ em x e $l-1$ em y. Se $f(a, b) = 0$ para todo o $a \in A$ e $b \in B$, então $f(x, y) \equiv 0$.*

Demonstração. Considere-se $f(x, y)$ um polinómio com coeficientes num corpo $\mathbb{F}$ e com grau no máximo $k-1$ em x e $l-1$ em y. Assim é possível escrever f como

$$f(x, y) = \sum_{i=0}^{k-1} \sum_{j=0}^{l-1} f_{i,j} x^i y^j,$$

ou, equivalentemente,

$$f(x, y) = \sum_{i=0}^{k-1} v_i(y) x^i$$

se se tomar

$$v_i(y) = \sum_{j=0}^{l-1} f_{i,j} y^j \text{ um polinómio de grau no máximo } l-1 \text{ em } y.$$

Para cada elemento $b \in B$ escolhido arbitrariamente, tem-se

$$g(x) = f(x, b) = \sum_{i=0}^{k-1} v_i(b) x^i.$$

O polinómio g é assim um polinómio de grau no máximo $k-1$ em x tal que $g(a) = 0\ \forall a \in A$. Contudo, $|A| = k$ e portanto g é um polinómio de grau no máximo $k-1$ e k raízes. Assim g apenas poderá ser o polinómio nulo. Consequentemente, deduz-se que $v_i(b) = 0\ \forall b \in B$. Usando novamente o argumento anterior, isto é, v_i é um polinómio de grau no máximo $l-1$ que tem pelo menos $l(= |B|)$ raízes, resulta que este polinómio também é o polinómio nulo.
Logo, $f_{i,j} = 0$ para quaisquer índices i, j e portanto $f(x, y) \equiv 0$. □

Lema 3.3.2. *Seja A um subconjunto finito de um corpo $\mathbb{F}$ e seja $|A| = k$. Para qualquer $m \geqslant 0$ existe um polinómio $r_m(x) \in \mathbb{F}[x]$ de grau no máximo $k-1$ tal que $r_m(a) = a^m$ para todo o elemento $a \in A$.*

Uma vez que esta secção se intitula de Método Polinomial é apenas natural que esta demonstração, à semelhança da anterior, envolva polinómios. A prova que se apresenta de seguida torna-se bastante elementar quando considerado o algoritmo da divisão para polinómios.

Demonstração. Seja t o polinómio de grau k com coeficientes em $\mathbb{F}[x]$ construido da seguinte forma:

$$t(x) = \prod_{j=0}^{k-1}(x - a_j).$$

Aplicando agora o algoritmo da divisão para polinómios sobre um corpo deduzir-se-á o pretendido.

Dito isto e tendo em conta o enunciado do Lema, o polinómio que se procura é, como esperado, o polinómio resto que resultará de uma conveniente divisão entre outros dois polinómios. Ora, visto que o polinómio t se anula sempre que é calculado para qualquer valor de A, o polinómio t fará o papel de polinómio divisor — de forma a anular-se e restar apenas o polinómio resto. Por outro lado, o polinómio dividendo terá de ser o polinómio x^m de forma a se conseguir obter o pretendido.

Efetuada então esta divisão obtém-se

$$x^m = t(x)q_m(x) + r_m(x)$$

e, portanto, quando aplicada a qualquer elemento a de A vem que

$$a^m = t(a)q_m(a) + r_m(a) = r_m(a).$$

Resta observar o grau do polinómio r. Contudo, pelo algoritmo da divisão garante-se que o grau de r é inferior ao grau de t logo $\deg(r) \leqslant k - 1$. □

Tendo em conta estes dois lemas, apresenta-se de seguida o teorema central deste artigo, a partir do qual é possível retirar uma demonstração para a conjetura de Erdős e Heilbronn.

Teorema 3.3.1. *Seja p um número primo e sejam A e B dois subconjuntos não vazios de $\mathbb{Z}_p$ tais que $|A| \neq |B|$. Seja*

$$C = \{a + b : a \in A, b \in B \text{ e } a \neq b\}.$$

Então

$$|C| \geqslant \min\{p, |A| + |B| - 2\}.$$

Demonstração. Sem perda de generalidade comece-se por assumir que
$|A| = k > |B| = l$, assim

$$1 \leqslant l < k \leqslant p.$$

Uma vez que o resultado dá um critério de minimalidade entre duas quantidades, esta demonstração será dividida em duas partes. Cada uma delas consiste em tomar o maior valor em consideração entre p e $|A| + |B| - 2$. Quando se supuser que um certo valor é maior que outro então o objetivo será ver que de facto a cardinalidade de C ainda é superior ao valor menor.

Suponha-se que $k + l - 2 > p$, pretende-se ver que $|C| > p$.

Veja-se que para $|B| = 1$ o resultado é válido. Pela hipótese acima considerada, virá que $k - 1 > p$. Assim, a cardinalidade da soma de elementos distintos é, pelo menos,$|A| - 1$, uma vez que pode existir no máximo uma repetição de elementos. Deste modo,

$$|C| \geqslant |A| - 1 = k - 1,$$

que, por hipótese, é maior que p.

A prova da hereditariedade consistirá em encontrar um conjunto que se irá denotar por C' com cardinalidade inferior à cardinalidade de C mas que verifique o resultado. Assim, sem outra possibilidade, C verificará o resultado também.
Como $k + l - 2 > p$, tem-se que $l > p - k + 2$ e escolha-se $l' := p - k + 2$. Assim

$$2 \leqslant l' < l < k \qquad \text{e} \qquad k + l' - 2 = p.$$

Considere-se então um conjunto B' contido em B tal que $|B'| = l'$ e seja

$$C' = \{a + b' : a \in A, b' \in B', a \neq b'\}.$$

Claramente temos $C' \subset C$. Ora, sendo o teorema válido, por indução forte, para os conjuntos A, B' e C', tem-se que

$$|C'| \geqslant \min\{p, |A| + |B'| - 2\} = k + l' - 2 = p = \min\{p, |A| + |B| - 2\}.$$

Como por outro lado, $|C| > |C'|$, sairá tal como pretendido

$$|C| > p.$$

Considere-se agora que $k + l - 2 < p$. Prove-se que se pode concluir que $|C| > k + l - 2$.
Seja a prova feita com recurso à técnica do absurdo. Suponha-se que $|C| < k + l - 2$, ou equivalentemente, que $|C| \leqslant k + l - 3$.

Seja $w \in \mathbb{Z}$ tal que $w + |C| = k + l - 3$. Construa-se o polinómio $f \in \mathbb{F}[x, y]$ do seguinte modo:

$$f(x, y) = (x - y)(x + y)^w \prod_{c \in C} (x + y - c).$$

Pela forma como foi construído f tem grau exactamente igual a $k+l-2$, quer na variável x como na variável y. E, além disto, $k+l-2$ é o grau total. Note-se que o produto $(x-y)(x+y)^w$ pode ser escrito como

$$(x-y)\sum_{j=0}^{w}\binom{w}{j}x^{w-j}y^j = \sum_{j=0}^{w}\binom{w}{j}x^{w-j+1}y^j - \sum_{j=0}^{w}\binom{w}{j}x^{w-j}y^{j+1}.$$

É então fácil de perceber que $(x-y)(x+y)^w$ se trata de um polinómio homogéneo com grau total $w+1$. Para obter f a partir do polinómio anterior, resta agora multiplicar por tantos fatores de grau 1 em x e em y quantos os elementos de C. Por conseguinte, $f(x,y)$ tem grau $w+1+|C|$ tanto em x como em y.
Sejam $a \in A$ e $b \in B$. Vejamos que $f(a,b) = 0$.
Se $a = b$, então:

$$f(a,b) = (a-b)(a+b)^w \prod_{c\in C}(a+b-c) = 0(a+b)^w \prod_{c\in C}(a+b-c) = 0.$$

Porém se $a \neq b$, então $a+b \in C$ e

$$f(a,b) = (a-b)(a+b)^w \prod_{c\in C}(a+b-c) = (a-b)(a+b)^w 0 = 0.$$

É então possível escrever $f(x,y)$ na seguinte forma:

$$\begin{aligned} f(x,y) &= \sum_{\substack{i,j\geqslant 0 \\ i+j\leqslant k+l-2}} f_{i,j}x^i y^j \\ &= (x-y)(x+y)^{k+l-3} + \text{ termos de grau inferior.} \end{aligned}$$

Visto que $1 \leqslant l < k \leqslant p$ (tal como visto no início) e $1 \leqslant k+l-3 \leqslant p-1$, vem que o coeficiente $f_{k-1,l-1}$ do monómio $x^{k-1}y^{l-1}$ em $f(x,y)$ é

$$\binom{k+l-3}{k-2} - \binom{k+l-3}{k-1}.$$

Primeiro observe-se que de facto este é o coeficiente certo. Ora, o monómio $x^{k-1}y^{l-1}$ é o monómio de maior grau e, pela última expressão apresentada de f conclui-se que este coeficiente apenas

aparecerá no produto $(x-y)(x+y)^{k+l-3}$. Sabe-se então que

$$\begin{aligned}(x-y)(x+y)^{k+l-3} &= (x-y)\sum_{j=0}^{k+l-3}\binom{k+l-3}{j}x^{k+l-3-j}y^j \\ &= \sum_{j=0}^{k+l-3}\binom{k+l-3}{j}x^{k+l-3-j+1}y^j - \sum_{j=0}^{k+l-3}\binom{k+l-3}{j}x^{k+l-3-j}y^{j+1} \\ &= \sum_{j=0}^{k+l-3}\binom{k+l-3}{j}x^{k+l-2-j}y^j - \sum_{j=0}^{k+l-3}\binom{k+l-3}{j}x^{k+l-3-j}y^{j+1}.\end{aligned}$$

Assim, o coeficiente pretendido ocorre como uma soma do termo em que $j = l-1$ no primeiro somatório com o termo em que $j = l-2$ no segundo somatório, logo

$$\begin{aligned}f_{i,j} &= \binom{k+l-3}{l-1} - \binom{k+l-3}{l-2} \\ &= \binom{k+l-3}{k+l-3-(l-1)} - \binom{k+l-3}{k+l-3-(l-2)} \\ &= \binom{k+l-3}{k-2} - \binom{k+l-3}{k-1}.\end{aligned} \tag{3.3}$$

Utilizando a definição de binómio de Newton e efetuando agora algumas simplificações verifica-se que

$$\binom{k+l-3}{k-2} - \binom{k+l-3}{k-1} = \frac{(k+l-3)!(k-l)}{(k-1)!(l-1)!}.$$

Note-se que este coeficiente não é divisível por p uma vez que tanto $(k+l-3)$ como $(k-l)$ são inferiores a p. Logo $f_{i,j}$ está de facto bem definido, tem a expressão acima mencionada e é não nulo módulo p. Desta forma confirma-se que o polinómio f tem grau $k+l-2$.

De seguida, procede-se a uma aplicação sucessiva do Lema 3.3.2 de forma a construir um novo polinómio à custa de f. Este novo polinómio, que será denotado por f^*, estará nas condições de aplicação do Lema 3.3.1 e, da sua conclusão será retirado o absurdo procurado.

Ora, o Lema 3.3.2 afirma que para qualquer $m \geqslant 0$ existe um polinómio $r_m(x)$ com grau no máximo $k-1$ tal que $r_m(a) = a^m$ para qualquer $a \in A$. Desta forma, também se tem que para qualquer $n \geqslant 0$ existe um polinómio $s_n(y)$ com grau no máximo $l-1$ tal que $s_n(b) = b^n$ para qualquer $b \in B$.

Recorde-se a seguinte expressão para f:

$$f(x,y) = \sum_{\substack{i,j\geqslant 0 \\ i+j\leqslant k+l-2}} f_{i,j}x^iy^j.$$

Assim,

$$\begin{aligned} f(x,y) = \sum_{\substack{i,j\geqslant 0 \\ i+j\leqslant k+l-2}} f_{i,j}x^iy^j = {} & f_{0,0}x^0y^0 + \ldots + f_{0,k+l-2}x^0y^{k+l-2} + \\ & + f_{1,0}x^1y^0 + \ldots + f_{1,k+l-3}x^1y^{k+l-3} + \\ & + \ldots + \\ & + f_{i,0}x^iy^0 + \ldots + f_{i,k+l-i-2}x^iy^{k+l-i-2} + \\ & + \ldots + \\ & + f_{k+l-1,0}x^{k+l-1}y^0 + f_{k+l-1,1}x^{k+l-1}y^1 \\ & + f_{k+l-2,0}x^{k+l-2}y^0. \end{aligned}$$

O argumento que se segue é semelhante ao que já tinha sido utilizado anteriormente quando se fez a separação entre expoentes "grandes" e expoentes "pequenos". Porém será agora necessário adaptar o argumento uma vez que neste caso f trata-se de uma função de duas variáveis.

Construa-se um novo polinómio, $f^*(x,y)$ de tal forma que:
se x^my^n for um monómio em $f(x,y)$ com $m \geqslant k$, isto é, de f tiver expoente "grande" na variável x, então substitui-se x^my^n por $r_m(x)y^n$. Tal como anteriormente, repare-se que estes casos são mutuamente exclusivos, isto é, expoente "grande" na variável x implica que f tenha expoente "pequeno" na variável y pois dado que $\deg(f) = k+l-2$ e que $m \geqslant k$, segue que $n \leqslant l-2$. Deste modo $r_m(x)y^n$ é uma soma de monómios x^iy^j com $i \leqslant k-1$ e $j \leqslant l-2$. Analogamente, se x^my^n for um monómio em $f(x,y)$, com $n \geqslant l$ substitui-se x^my^n por $x^ms_n(y)$.
Com estas substituições obtém-se o polinómio $f^*(x,y)$ que tem grau no máximo $k-1$ em x e $l-1$ em y.

Repare-se ainda que $f^*(a,b) = 0$ para cada $a \in A$ e $b \in B$ pois $f^*(a,b) = f(a,b) = 0$.
Então, aplicando o Lema 3.3.1, vem que $f^*(x,y) \equiv 0$, o que é absurdo pois tinha ficado provado que o coeficiente $f_{k-1,l-1}$ era não nulo em f e, consequentemente, em f^*. Conclui-se assim a demonstração do teorema. □

Com o auxílio desta nova técnica polinomial, a prova da famosa conjetura sai como corolário, vejamos:

Teorema 3.3.2. *Seja p um número primo e seja A um subconjunto de $\mathbb{Z}_p$ tal que $|A| = k \geqslant 2$. Denote-se por $2A$ o conjunto de todas as somas de dois elementos distintos de A. Então*

$$|2A| \geqslant \min\{p, 2k-3\}.$$

Demonstração. Seja B o subconjunto de $\mathbb{Z}_p$ tal que $B = A \setminus \{a\}$ para algum $a \in A$. Desta forma $|B| = |A| - 1$. Seja C tal como definido no teorema anterior, $C = \{a+b : a \in A, b \in B, a \neq b\}$.

Obtém-se que $C \subset 2A$ logo, pelo teorema anterior resulta que

$$|2A| \geqslant |C| \geqslant \min\{p, |A| + |B| - 2\} = \min\{p, 2|A| - 3\}.$$

□

Para além de uma demonstração alternativa para a conjetura de Erdős e Heilbronn, o método polinomial permite também demonstrar o Teorema de Cauchy-Davenport seguindo uma demonstração bastante semelhante à que foi apresentada para o Teorema 3.3.1. Este método pode também ser aplicado na demonstração do seguinte resultado, a qual omitimos deste texto mas pode ser encontrada em [1].

Teorema 3.3.3. *Seja p um número primo e sejam A e B subconjuntos não vazios de $\mathbb{Z}_p$, tais que $|A| = k$ e $|B| = l$. Considere-se $C = \{a + b : a \in A, b \in B \text{ e } ab \neq 1\}$. Então*

$$|C| \geqslant \min\{p, |A| + |B| - 3\}.$$

4. Conclusão

Todo o desenvolvimento anteriormente apresentado foi possível devido ao enorme trabalho realizado pelos professores Dias Da Silva e Hamidoune, comprovando que, por vezes, a resposta para algumas questões em aberto poderá já estar ao nosso alcance, "apenas" teremos de conseguir decifrar a ligação entre as duas. É de tamanha importância, na minha opinião, podermos presenciar estes avanços e, tal como se tem revelado até aos dias de hoje, este tema permitirá continuar a investigação noutras áreas de Álgebra, sendo que este legado continua vivo através das investigações do matemático Terence Tao, entre outros.

Por fim, gostaria de reforçar que a ideia introdutória não pretende de nenhum modo desencorajar a procura de novas teorias originais, apenas considero que seja igualmente importante olhar para o que já foi alcançado até agora e estudar as suas possíveis aplicações noutras áreas, conduzindo à possibilidade de, tal como aconteceu aqui, provar de conjeturas com várias décadas.

Não obstante, o importante será sempre o enriquecimento da Matemática, quer seja através do desenvolvimento de novas teorias quer seja através do estabelecimento de ligação entre áreas aparentemente distantes e, para o qual, devemos estar sempre despertos e recetivos.

Bibliografia

[1] N. Alon, M. B. Nathanson, and I. Rusza. Adding Distinct Congruence Classes Modulo a Prime. *American Mathematical Monthly*, 102(3):250–255, 1995.

[2] R. Bhatia. *Matrix Analysis*. Springer-Verlag New York, 1997.

[3] Cauchy. Recherches sur les nombres. *Journal de l'École polytechnique*, XVI(IX):39–63, 1813.

[4] H. Davenport. On the Addition of Residue Classes. *Journal of the London Mathematical Society*, pages 30–32, 1935.

[5] H. Davenport. A Historical Note. *Journal of the London Mathematical Society*, (22):100–101, 1947.

[6] J. A. Dias da Silva and Y. O. Hamidoune. A Note on the Minimal Polynomial of the Kronecker Sum of Two Linear Operators. *Linear Algebra and its Applications*, (141):283–287, 1990.

[7] J. A. Dias da Silva and Y. O. Hamidoune. Cyclic Subspaces for Grassmann Derivatives and Additive Theory. *The Bulletin of the London Mathematical Society*, (26):140–146, 1994.

[8] F. R. Gantmacher. *The Theory of Matrices, Vol I.* Chelsea Publishing Company, 1960.

[9] S. Guo and Zhi-Wei Sun. A variant of Tao's method with application to restricted sumsets. *Journal of Number Theory*, 129(2):434–438, 2009.

[10] H. Henderson, F. Pukelsheim, and S. R. Searle. On the History of the Kronecker Product. *Linear and Multilinear Algebra*, 14(2):113–120, 1983.

[11] L. Hogben. *Handbook of Linear Algebra, Part I.* Chapman and Hall/CRC, 2013.

[12] R. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

[13] S. Lang. *Algebra*. Addison-Wesley, Reading, MA, 1972.

[14] M. Marcus. *Finite Dimensional Multilinear Algebra, Part I.* M. Dekker, 1973.

[15] R. Merris. *Multilinear Algebra*. CRC Press, 1997.

[16] M. B. Nathanson. *Additive Number Theory - Inverse Problems and the Geometry of Sumsets.* Springer, 1996.

[17] T. Tao. An Uncertainty Principle for Cyclic Groups of Prime Order . *Mathematical Research Letters*, 12(1):121–127, 2005.